numero 70
giugno 1974
anno VII
volume XII

LE SCIENZE
edizione italiana di SCIENTIFIC AMERICAN

La gassificazione del carbone

Questa tecnologia, soppiantata dall'impiego del gas naturale e del petrolio, suscita attualmente un rinnovato interesse. Nuovi metodi offrono fondate promesse per una sorgente alternativa di energia da combustibile fossile

di Harry Perry

Quasi tutte le maggiori città dell'est degli Stati Uniti avevano un tempo la propria fabbrica del gas, in cui quest'ultimo veniva prodotto (generalmente a partire dal carbone) per l'illuminazione e per gli usi di cucina. Queste fabbriche e i massicci serbatoi cilindrici che sorgevano nelle loro vicinanze scomparvero gradualmente dopo la seconda guerra mondiale quando il gas naturale cominciò a essere distribuito su scala nazionale attraverso le reti di gasdotti. Ma ora, con il gas naturale e il petrolio che scarseggiano sempre di piú, l'idea di migliorare la situazione energetica nazionale per mezzo della gassificazione del carbone, che al contrario è ancora abbondante, sta polarizzando un notevole interesse.

Il carbone è stato trovato in trenta dei cinquanta stati degli Stati Uniti e rappresenta piú del 90 per cento delle riserve accertate di tutti i combustibili finora noti. Con una rapidità di consumo pari a quella registrata nel 1972 le riserve accertate di carbone non dovrebbero esaurirsi per circa 600 anni. Anche se il carbone restasse l'unica sorgente di energia e la richiesta totale di quest'ultima aumentasse in ragione del 3,5 per cento all'anno, le riserve accertate durerebbero per 47 anni mentre il totale delle riserve di carbone per quasi 75 anni.

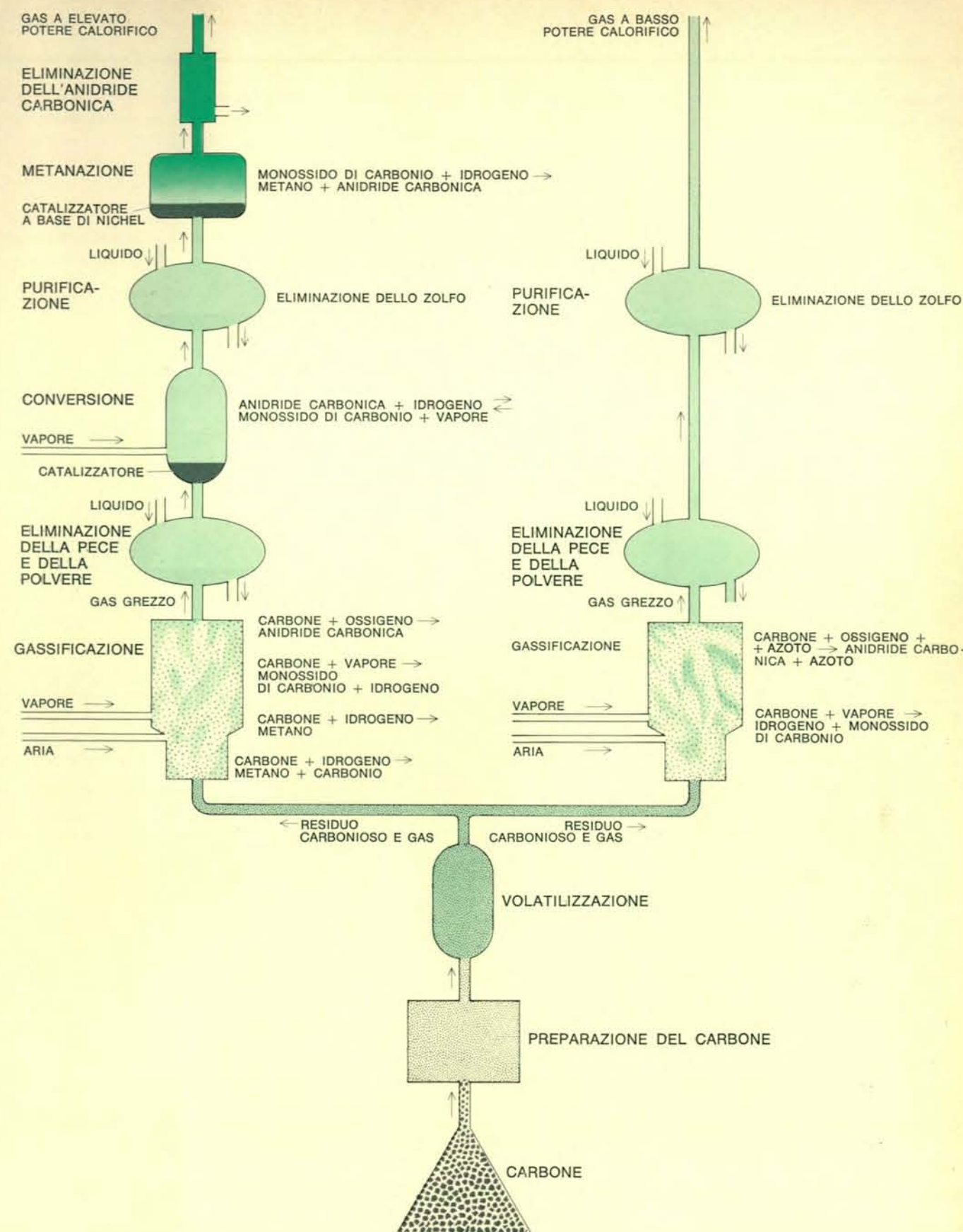
Il gas naturale è giunto al proprio dominio per il fatto che il suo costo è basso mentre il suo potere calorifico è elevato. Allo stesso modo di altri gas esso è pulito, facile da distribuire e insomma di uso piú che conveniente. Negli anni seguenti alla seconda guerra mondiale esso divenne per questi motivi il combustibile preferito per gli usi domestici e commerciali, trovando nel contempo anche diverse applicazioni industriali. Per questi motivi il consumo di gas naturale è cresciuto in ragione del 5,4 per cento all'anno fra il 1947 e il 1971, in confronto al 3,1 per cento all'anno relativo al consumo totale di energia.

Il gas prodotto dal carbone potrebbe sostituire il gas naturale, sebbene le vecchie tecnologie di gassificazione del carbone dovrebbero essere migliorate allo scopo di fornire al prodotto lo stesso potere calorifico del gas naturale (circa 9200 calorie al metro cubo). Inoltre il gas prodotto dal carbone e non arricchito (dotato quindi di un potere calorifico compreso fra le 1100 e le 5000 calorie al metro cubo) potrebbe servire egregiamente in una grande quantità di applicazioni industriali. Pertanto non è affatto sorprendente il fatto che la gassificazione del carbone venga ora riesaminata come sorgente di primaria importanza per un'energia pulita adatta agli usi domestici.

Il gas prodotto dal carbone è stato ottenuto per la prima volta nel diciottesimo secolo riscaldando il carbone in assenza di aria. La prima compagnia per il gas di carbone, la quale distribuiva il suo prodotto per l'illuminazione, venne fondata a Londra nel 1812; la prima negli Stati Uniti venne fondata a Baltimora nel 1816.

A quei tempi il gas veniva prodotto attraverso una distillazione distruttiva, ovvero riscaldando il carbone fino a una temperatura alla quale esso si decomponeva chimicamente. Il gas prodotto con la distillazione del carbone aveva un potere calorifico compreso fra le 4200 e le 5000 calorie al metro cubo, a seconda del tipo di carbone e della temperatura alla quale esso veniva riscaldato. Un tipo di gas del tutto simile lo si ottiene quando si procede alla cosiddetta distillazione secca o « carbonizzazione » del carbone per ottenere coke; assai spesso il gas proveniente dai forni di calcinazione per coke ha sostituito il gas di carbone quando e dove il primo era facilmente disponibile mentre il secondo scarseggiava.

Nel corso di questi processi il 70 per cento o piú del carbone originale resta sotto forma di un residuo solido, costituito principalmente di carbone impoverito di sostanze volatili che deve essere venduto o utilizzato altrimenti. Nel caso dei forni per coke questo fatto non costituisce un problema, poiché il coke è il prodotto fondamentale ottenuto e il gas può considerarsi un sottoprodotto.

La gassificazione implica non solo il riscaldamento del carbone, come nel caso della distillazione, ma anche la successiva reazione del residuo solido con aria, ossigeno, vapore oppure diverse miscele di essi. Dalla fase di distillazione si ottiene una certa quantità di gas dotato di un potere calorifico discretamente elevato per il fatto che il metano (CH_4) e gli altri idrocarburi di ordine piú elevato contenuti nel carbone sono praticamente i primi componenti che emergono dalla decomposizione del carbone. Dalla fase di gassificazione, invece, si ottiene un gas costituito essenzialmente da una miscela di

Le diverse fasi del processo di gassificazione del carbone sono qui rappresentate per due tipi di gas (*in colore*). In entrambi i casi il carbone viene preparato secondo le necessità e quindi viene volatilizzato, operazione questa che comporta il suo riscaldamento in assenza di aria in modo da ottenere una decomposizione chimica, da cui si sviluppano vari tipi di gas insieme a un residuo carbonioso. Se l'obiettivo che ci si propone con tale processo è quello di ottenere un gas con un potere calorifico paragonabile a quello del gas naturale, le fasi successive alla volatilizzazione sono quelle illustrate a sinistra della figura. Il processo relativamente piú semplice schematizzato a destra consente di ottenere un gas di gassogeno, il quale è caratterizzato da un potere calorifico molto piú basso ma si presta bene a essere impiegato come combustibile per caldaia.

idrogeno e monossido di carbonio (quest'ultimo è caratterizzato da un potere calorifico molto piú basso di quello proprio dei gas di distillazione) insieme ad alcuni dei gas distillati dal carbone. La quantità del gas di distillazione presente nel prodotto finale dipende dal processo di gassificazione usato.

L'ossigeno è alquanto costoso, e cosí le vecchie compagnie del gas si rivolsero a un processo alternativo per produrre un gas da poter aggiungere al gas di carbone e a quello dei forni per coke. Il processo produceva una sostanza conosciuta con il nome di « gas blu » oppure di « gas d'acqua ». Del carbone del tipo antracite, da cui venivano preventivamente asportati i pezzi piú piccoli, mentre quelli piú grossi venivano ridotti a dimensioni accettabili, era introdotto in un recipiente rivestito con un materiale refrattario. Non appena il gassificatore raggiungeva una temperatura sufficientemente elevata, il carbone (ma poteva anche essere il coke) veniva lambito da una corrente d'aria e riscaldato ulteriormente dalla combustione risultante con l'ossigeno presente nell'aria. Durante le prime fasi di questo periodo di « soffiatura », il gas prodotto risultava prevalentemente anidride carbonica ($C+O_2 \rightarrow CO_2$), ma quando il carbone o il coke degli strati superiori della massa diventava ancora piú caldo cominciava a innescare la reazione inversa ($CO_2+C \rightarrow CO$), e questo significa che il monossido di carbonio veniva prodotto in quantità sempre piú grandi. Quando una sufficiente quantità di carbone o di coke aveva raggiunto una temperatura abbastanza elevata, le valvole che controllavano il flusso dell'aria venivano chiuse mentre altre valvole venivano aperte per iniziare la « fase attiva », durante la quale il carbone o il coke incandescente veniva fatto reagire con il vapor d'acqua ($C+H_2O \rightarrow H_2+CO$). Queste fasi dovevano susseguirsi in un processo ciclico, e tale fatto portava inevitabilmente a notevoli complicazioni nel progetto e nel funzionamento a regime dell'impianto.

Il gas d'acqua aveva un potere calorifico di circa 2700 calorie al metro cubo, in confronto a quello del gas di carbone compreso fra le 4200 e le 5000 calorie al metro cubo, e pertanto esso doveva essere arricchito. Tale arricchimento implicava un altro processo ciclico. Il gas ottenuto alla fine di una fase attiva era caldissimo e conteneva abbastanza monossido di carbonio da poter bruciare a sua volta e riscaldare cosí dei mattoni refrattari contenuti in un recipiente detto « carburatore ». Quando i mattoni erano caldi abbastanza, veniva interrotto il flusso del gas mentre si spruzzava dell'olio sui mattoni; qui l'intenso calore provocava il cracking dell'olio, che si trasformava in idrocarburi assai piú leggeri come per esempio il metano e il propano. Il gas che si otteneva in tal modo, caratterizzato da un elevato potere calorifico, veniva mescolato al gas ottenuto durante la fase attiva del processo per dare origine a quello conosciuto con il nome di « gas blu carburato ».

Da principio le compagnie vendevano un gas di carbone dotato di un potere calorifico di circa 4900 calorie al metro cubo. I fornelli venivano pertanto progettati per funzionare con un gas di tali caratteristiche. Una volta determinato il progetto dei bruciatori, il gas fabbricato doveva essere comunque adattato a quel preciso potere calorifico, indipendentemente dal processo di fabbricazione seguito. Quando si introdusse l'uso del gas naturale, in molte zone divenne necessario produrre un gas che potesse essere miscelato con quello naturale. Per fortuna il potere calorifico del gas carburato poteva essere controllato su un ampio spettro di valori semplicemente variando il tipo e la quantità dell'olio spruzzato sui mattoni refrattari.

Esisteva anche il « gas di gassogeno », che era ancora piú economico del gas d'acqua carburato. Esso era caratterizzato da un potere calorifico variabile fra le 1000 e le 1450 calorie al metro cubo, ciò che implicava che esso non poteva essere impiegato per un servizio di distribuzione perché il prezzo sarebbe salito a livelli inaccettabili. (I costi di distribuzione incidono per la maggior parte della bolletta degli utenti; anche con il gas naturale i costi di distribuzione nel Distretto di Columbia ammontano al 65 per cento del prezzo totale al consumo.) Il gas di gassogeno serviva egregiamente, tuttavia, nelle industrie che avevano bisogno di un combustibile pulito da impiegare per la combustione o come sorgente di calore in diversi processi di lavorazione e che, nel contempo, potevano fabbricare il gas sul posto, in modo che i costi di distribuzione risultassero sufficientemente contenuti.

Il gas di gassogeno veniva realizzato con un procedimento di natura continua, secondo il quale una massa di carbone o di coke incandescente veniva lambita da una corrente d'aria o da una miscela di aria e di vapor acqueo. Il prodotto finale conteneva necessariamente l'azoto derivato dall'aria e l'anidride carbonica risultante dalla combustione del carbone con l'ossigeno. È vero che l'anidride carbonica poteva essere separata, ma l'azoto diluiva in ogni caso il potere calorifico del gas. Questo non era un grave svantaggio quando il gas veniva impiegato come combustibile industriale, ma rendeva impossibile la sua utilizzazione come gas di sintesi (idrogeno e miscele di idrogeno e monossido di carbonio impiegate nella fabbricazione delle materie plastiche, dei combustibili e di altri prodotti sintetici) e per la distribuzione a scopi commerciali.

A partire dal 1850 circa e in progressione fino alla seconda guerra mondiale, la tecnologia di fabbricazione del gas d'acqua e del gas di gassogeno venne continuamente migliorata negli Stati Uniti. Dopo che il gas artificiale ebbe perduto la sua importanza commerciale negli Stati Uniti, tale tecnologia venne ulteriormente migliorata in Europa, dato che allora il carbone era l'unico combustibile che vi si trovasse in quantitativi significanti. Si studiarono cosí nuovi processi, sfruttando i vantaggi derivanti dai progressi tecnici conseguiti in altri campi come lo sviluppo di impianti per la produzione dell'ossigeno su larga scala, i nuovi metodi per il trattamento dei solidi reagenti con i gas e i miglioramenti ottenuti nei materiali da costruzione. Ma proprio nel momento in cui la tecnologia aveva raggiunto un livello che poteva giustificare l'investimento degli impianti industriali, vennero scoperti i giacimenti di gas naturale nel Mare del Nord e in Nord Africa. Per di piú, la maggior parte dei paesi europei operò una scelta tendente a dirigere l'economia dall'impiego del carbone estratto sul posto a costi elevati verso lo sfruttamento del petrolio che a quei tempi veniva importato a prezzi di assoluta convenienza. Cosí si costruirono pochissimi impianti di gassificazione del carbone impostati secondo le nuove tecnologie, mentre l'interesse per ulteriori miglioramenti delle stesse a poco a poco svanIva.

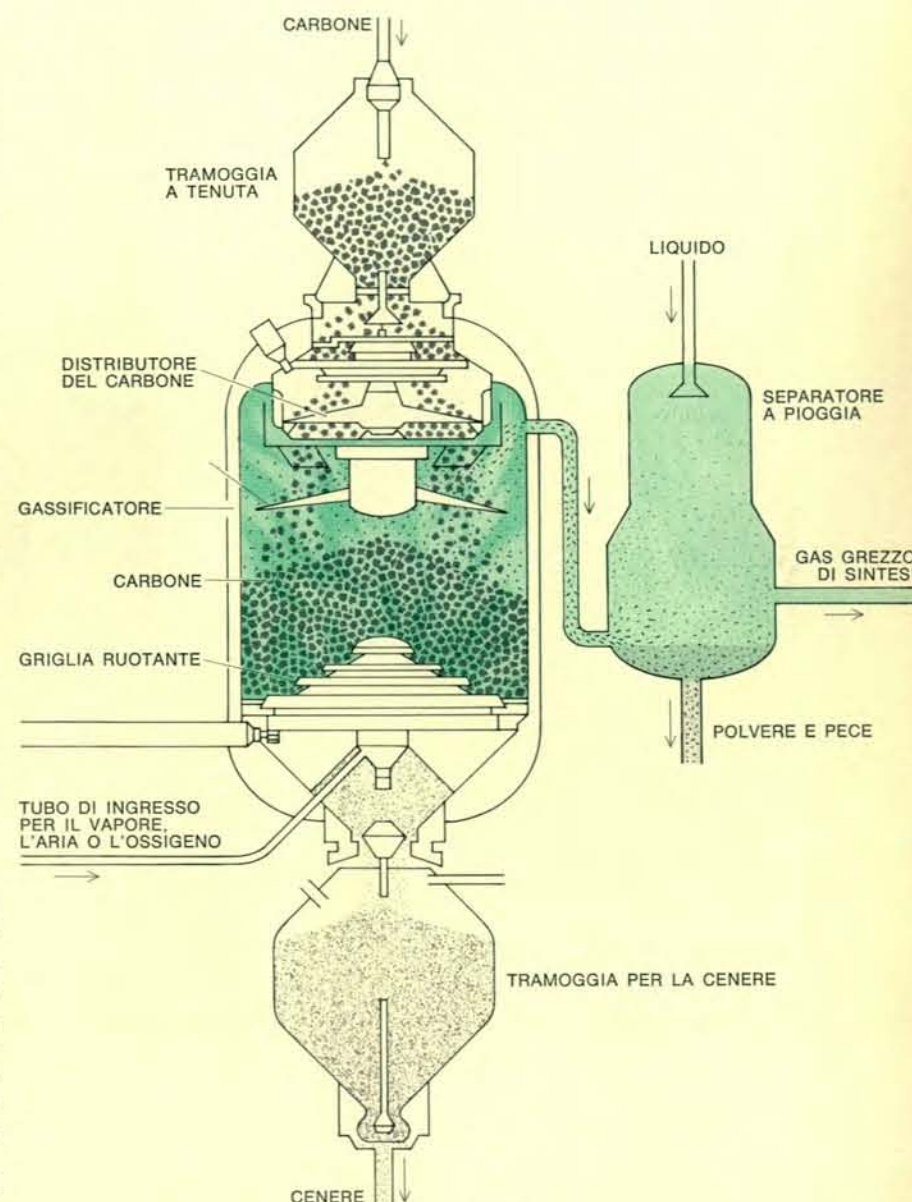
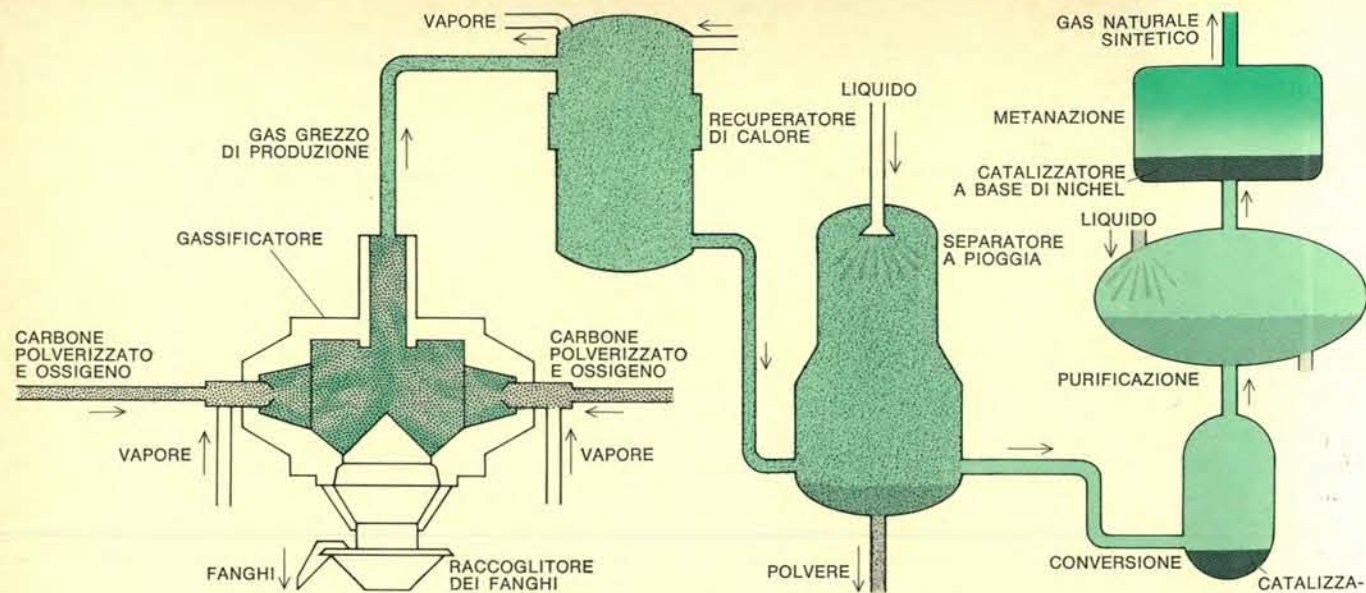
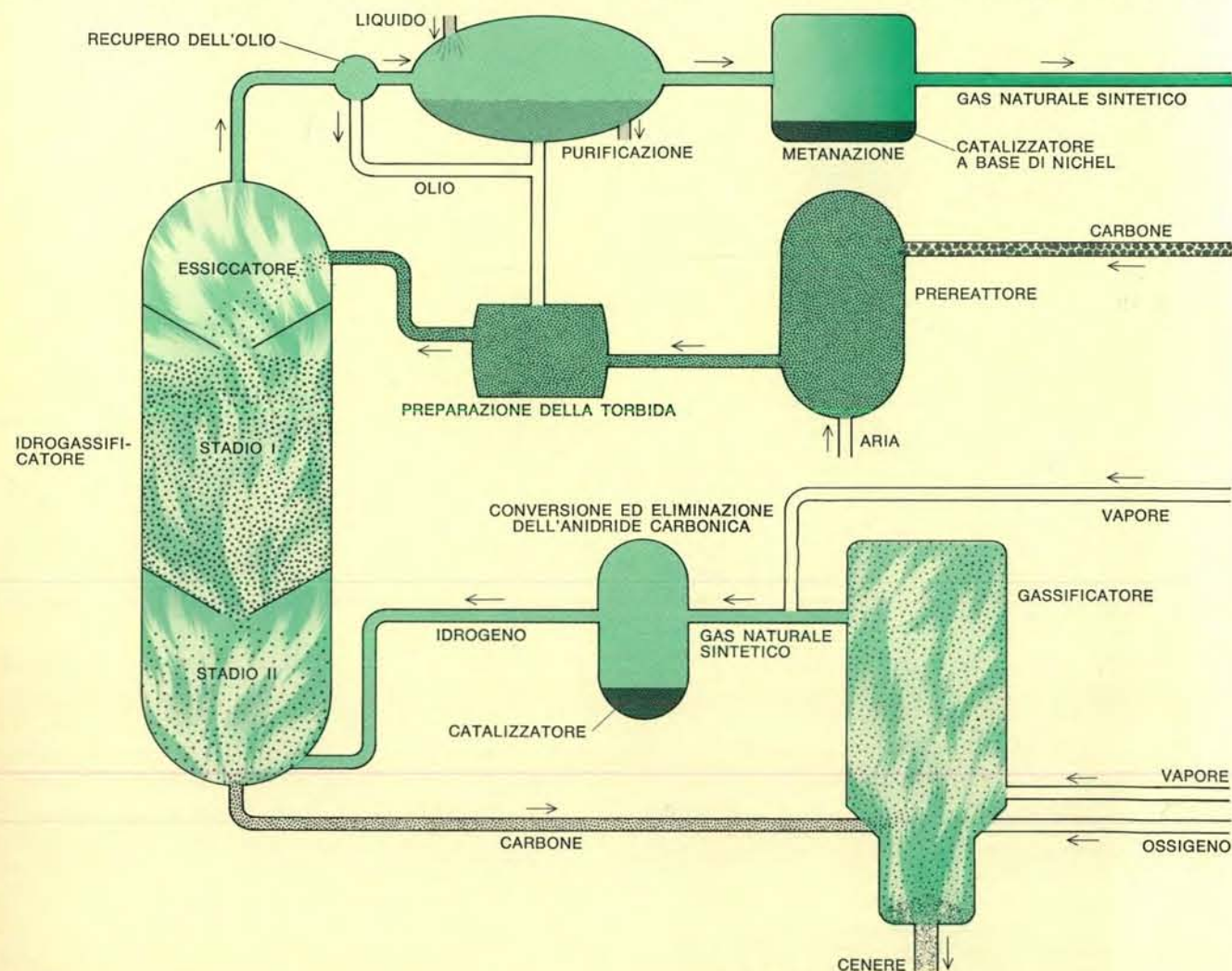
I due processi che furono realizzati piú spesso in Europa e altrove sono quello del gassificatore pressurizzato Lurgi e quello Koppers-Totzek a pressione atmosferica. Entrambi venivano impiegati nella maggioranza dei casi per fabbricare gas di sintesi, utilizzato a sua volta per la produzione di ammoniaca e di altri composti sintetici, sebbene, in qualche posto, impianti del tipo Lurgi provvedessero anche a distribuire il gas come gas di città. Nessuno di questi due procedimenti raggiungeva tuttavia tutte le specifiche qualitative attribuite a una reazione di gassificazione ideale. Un processo ideale dovrebbe essere costituito da un'unica operazione continua, che possa impiegare l'aria in qualità di componente di ossidazione; esso dovrebbe convertire qualunque tipo di carbone in un gas combustibile o in un gas di sintesi che abbia comunque un basso contenuto di costituenti inerti.

Sebbene sia il processo Lurgi sia il Koppers-Totzek siano a un solo stadio e continui, entrambi si basano sull'ossigeno piuttosto che sull'aria come agente di ossidazione. Il gassificatore di Lurgi è del tipo pressurizzato, fatto questo che per molte applicazioni del gas di sintesi porta a un vantaggio economico rispetto ai processi operanti a pressione atmosferica. D'altra parte il processo Lurgi richiede un carbone dalle dimensioni standardizzate (si deve eliminare quello piú fine) e di qualità non calcinata o tuttalpiú solo leggermente calcinata (perché se lo strato di carbone si calcina durante il processo di gassificazione, esso forma una massa compatta che impedisce il passaggio del gas attraverso lo strato medesimo fermando cosí il progredire dell'azione di gassificazione), mentre al contrario il processo Koppers-Totzek può funzionare con qualsiasi tipo di carbone.

Negli Stati Uniti il rinnovato interesse nella gassificazione del carbone per produrre un sostituto del gas naturale ha indotto a serie considerazioni sull'opportunità di installare degli impianti di tipo Lurgi o Koppers-Totzek. Nessuno dei due processi, tuttavia, è in grado di fornire un gas che possa sostituire quello naturale perché esso non ha il potere calorifico né la composizione richiesta. Il gas ottenuto da entrambi i processi è fondamentalmente costituito da una miscela di monossido di carbonio e di idrogeno, piú una certa quantità di metano distillato dal carbone durante la fase di riscaldamento. Il monossido di carbonio e l'idrogeno presentano entrambi un potere calorifico pari a 2700 calorie al metro cubo, molto al di sotto del potere calorifico del gas naturale. Per di piú, in certe zone, le compagnie del gas potrebbero anche non riuscire a ottenere il permesso di distribuire miscele di questo tipo (anche se una volta tali permessi venivano accordati molto piú facilmente) perché il monossido di carbonio è velenoso mentre l'idrogeno presenta dei problemi di contenimento e inoltre la sua combustione è del tutto particolare e richiede dei bruciatori speciali appositamente concepiti.

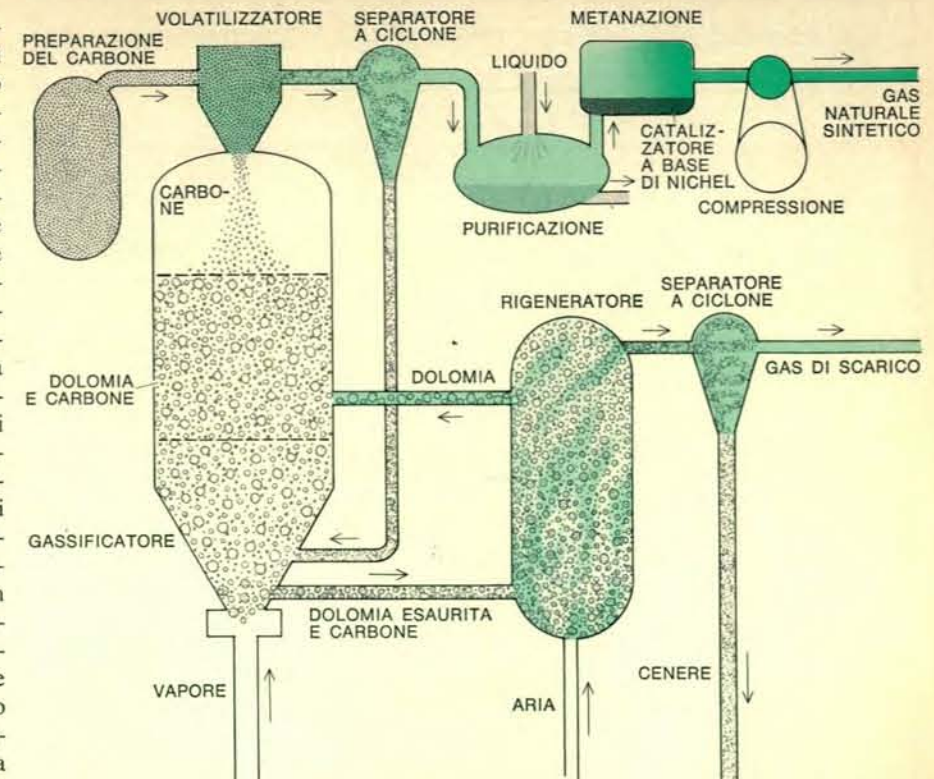
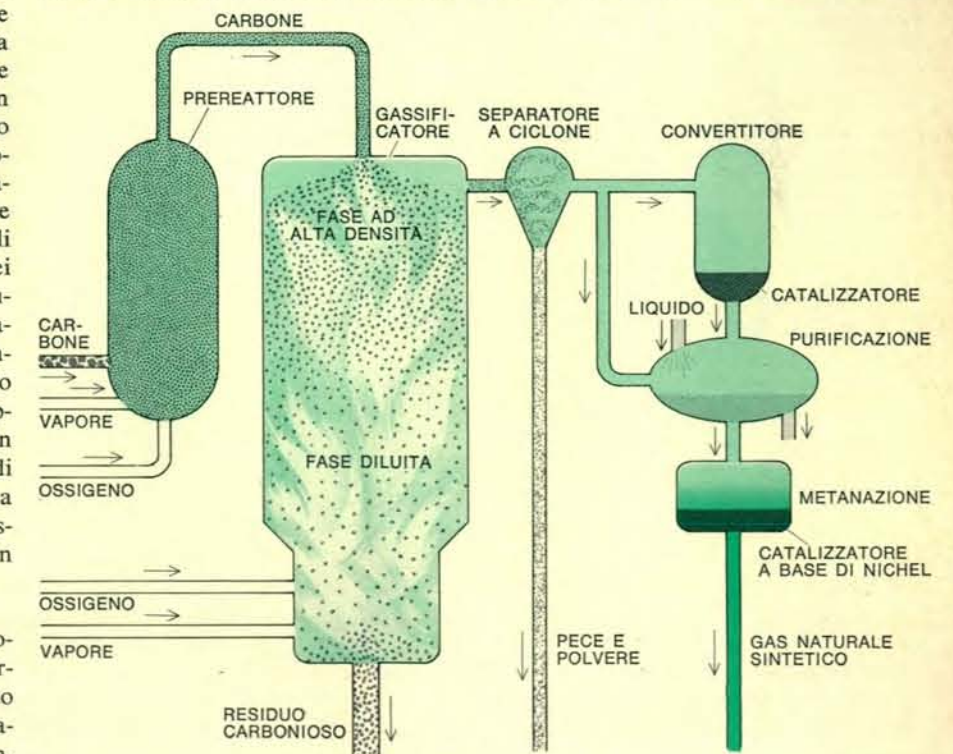
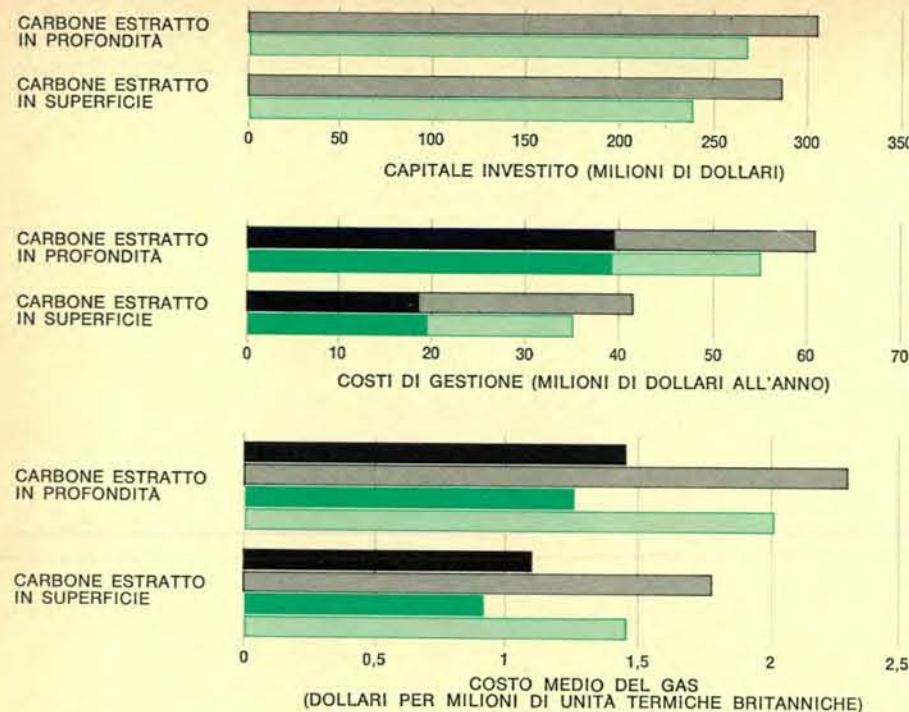
Entrambi i processi potrebbero tuttavia fornire un gas di sintesi grezzo che (in seguito a opportuna purificazione) potrebbe essere metanato in modo da ottenere un gas utilizzabile come sostituto del gas naturale. La metanazione implica il passaggio del gas su di un apposito catalizzatore al nichel allo scopo di ottenere la sua conversione in metano quasi completamente puro, che poi non è altro che il gas naturale. Ma purtroppo non è stato ancora trovato un procedimento di metanazione cosí efficiente per uso commerciale.

Sebbene il programma di ricerca negli Stati Uniti sulla gassificazione del carbone sia accentrato attualmente sulla realizzazione di un gas di sintesi economico dal quale sia possibile ricavare un sostituto del gas naturale, in un prossimo futuro la gassificazione del carbone dovrebbe essere impiegata anche per ottenere la materia prima per la produzione di altri prodotti sintetici (alcooli, chetoni, cere e altri derivati del petrolio) e dell'idrogeno necessario per ottenere i carburanti liquidi sintetici derivati dal carbone. (Il rapporto carbonio-idrogeno risulta molto piú elevato nel carbone che non nell'olio combustibile; la liquefazione del carbone contempla la produzione di idrogeno da parte del carbone con un processo di gassificazione nonché la reazione del gas con il carbone stesso allo scopo di incrementare il contenuto di idrogeno in modo da poter ottenere un olio combustibile.)

Inoltre le nuove leggi riguardanti le emissioni di ossidi di zolfo e le difficol-

Il processo Lurgi è uno dei due metodi per la gassificazione del carbone già realizzati su scala commerciale. È un processo che si svolge sotto pressione e in cui il carbone, opportunamente ridotto in pezzi, cade all'interno del gassificatore e viene dapprima essiccato e quindi carbonizzato in seguito alla reazione che si sviluppa con l'ossigeno e il vapore. Nello strato alla base del gassificatore il carbone rimasto viene fatto bruciare per fornire il calore necessario alle reazioni che si svolgono piú in alto.

Il processo Koppers-Totzek è il secondo metodo commercialmente valido per la gassificazione del carbone. Anche esso impiega vapore e ossigeno nella fase di gassificazione. Il processo si svolge a pressione atmosferica e può funzionare con qualsiasi tipo di carbone. Nessuno di questi processi commerciali è in grado di fornire un gas con un potere calorifico uguale a quello del gas naturale, per cui entrambi richiedono di essere seguiti da una fase di metanazione, in cui il gas prodotto dal carbone viene fatto passare su di un catalizzatore a base di nichel. Non si è ancora riusciti a ottenere un processo di metanazione che sia commercialmente valido, ma si stanno conducendo diversi esperimenti in tal senso.

Il processo HYGAS che viene sviluppato dall'Institute of Gas Technology è attualmente allo stadio di impianto pilota. Parte del carbone iniettato nel gassificatore viene gassificato in modo da produrre una miscela di idrogeno e di monossido di carbonio, la quale in seguito a una conversione si trasforma in idrogeno puro. Si fa quindi reagire questo idrogeno con il carbone, per ottenere un prodotto costituito fondamentalmente da metano. Il gas è poi sottoposto a una reazione di metanazione finale per elevarne la qualità. Il processo di gassificazione HYGAS si svolge ad alta pressione, su una massa di carbone fluidificato.

tà che si sono incontrate nel realizzare i procedimenti di purificazione dei fumi delle ciminiere hanno condotto necessariamente a un rinnovato interesse per la produzione di un gas a basso contenuto calorico e a basso tenore di zolfo che si possa ottenere dal carbone e impiegare come combustibile nelle caldaie. La maggior parte delle prime ricerche e dei procedimenti tentati per controllare le emissioni di ossidi di zolfo era imperniata sulla rimozione dello zolfo dai gas di ciminiera risultanti dalla combustione del carbone. Si è constatato che tutti i processi di purificazione dei fumi sinora sperimentati risultano non sufficientemente fidati per le elevate prestazioni che si richiedono nel caso delle caldaie private e industriali, e che i costi si rivelano molto piú elevati di quanto in un primo tempo si era sperato. Due problemi particolarmente gravi sono costituiti dall'enorme quantità di gas che deve essere sottoposto al trattamento (fatto questo che rende elevato il costo delle opere di intubamento) e da una eliminazione dei rifiuti solidi in qualche modo accettabile dal punto di vista ecologico.

Se si gassifica dapprima il carbone con l'aria in modo da ottenere un gas a basso contenuto calorico, il solfuro di idrogeno che ne risulta può essere eliminato con metodi già ampiamente sperimentati, prima che il gas venga fatto bruciare nella caldaia. Il volume di gas grezzo da trattare sarebbe in questo caso meno della metà di quello cui ci si trova di fronte quando si vuole asportare l'ossido di zolfo direttamente dal gas di ciminiera, e inoltre non si avrebbero fastidi ambientali per quanto riguarda l'eliminazione dei sottoprodotti. L'interesse nella produzione di un gas a basso contenuto calorico come mezzo per ottenere il controllo delle emissioni di ossidi di zolfo si è accresciuto ulteriormente non appena intravista la possibilità che gran parte dei processi avanzati in fase di sviluppo per la produzione di gas a elevato contenuto calorico potrebbe essere modificato in modo da ottenere un gas a basso contenuto calorico.

Frattanto sono allo studio molti nuovi metodi di gassificazione del carbone, e due di essi hanno raggiunto una prima fase realizzativa. Poiché tali ricerche si sono dirette quasi sempre verso la produzione di un sostituto del gas naturale, anche i procedimenti elaborati risultano orientati verso il tentativo di trattenere nel prodotto finale la maggiore quantità possibile del metano che si sviluppa durante la prima fase del processo, quando cioè il

Il processo « accettore » di anidride carbonica, un progetto messo a punto dalla Consolidation Coal Company, è anch'esso allo stadio di impianto pilota avanzato. Della dolomia calcinata viene fatta circolare attraverso una massa fluidificata di carbone tipo lignite. La dolomia reagisce con l'anidride carbonica prodotta durante la gassificazione iniziale del carbone, liberando cosí una quantità di calore sufficiente a sostenere la reazione carbone-vapore. Il gas grezzo prodotto, che contiene metano, idrogeno e monossido di carbonio, viene successivamente sottoposto a una fase finale di metanazione.

Il processo Synthane dell'US Bureau of Mines è attualmente allo stadio di impianto pilota sperimentale. Questo processo si prefigge lo scopo di produrre un gas di sintesi che possa servire come sostituto del gas naturale. Il carbone viene preventivamente trattato allo scopo di eliminare le sue proprietà agglomeranti, carbonizzato nella zona del gassificatore ad alta densità, e quindi viene gassificato con l'aiuto dell'ossigeno e del vapore nella zona a bassa densità. Si ha quindi una conversione che ha lo scopo di stabilizzare sul valore di 3:1 il rapporto fra idrogeno e monossido di carbonio, seguita da una reazione di metanazione ottenuta grazie a un catalizzatore a base di nichel.

I costi della gassificazione sono qui stimati nel caso di un processo Lurgi (*in grigio*) e nel caso di uno dei nuovi processi ora in fase sperimentale (*in colore*). Le parti piú scure delle strisce dei costi di gestione rappresentano il costo del carbone, e le parti piú chiare tutte le altre spese. Il costo medio del gas è stimato su di un periodo di 20 anni a partire dal 1975 (*strisce piú scure*) oppure dal 1990 (*strisce piú chiare*).

carbone viene semplicemente riscaldato. In questo modo si riduce il costo del capitale complessivamente impiegato per unità di metano prodotto, nonché i requisiti richiesti per i materiali. L'idea di trattenere il metano risulta pure vantaggiosa nella fabbricazione di un gas da caldaia a basso contenuto calorico, per il fatto che il metano contribuisce ad abbassare il prezzo e a elevare la qualità del prodotto.

I passi fondamentali richiesti per la produzione di un gas sintetico simile a quello naturale e ricavato dal carbone possono essere descritti in termini generali (*si veda la figura a pagina 9*). Dapprima il carbone viene preparato, operazione che comprende il taglio alle dimensioni massime desiderate, la rimozione dei frammenti piú piccoli (se necessaria) e dei materiali estranei, nonché la asciugatura (se risulta anch'essa necessaria per le fasi successive). Quindi il carbone viene inviato al gassificatore oppure, se è un carbone di tipo calcinato e il procedimento usato non può funzionare con tale carbone, esso subisce un pretrattamento a elevata temperatura allo scopo di eliminare le sue proprietà peculiari.

Il carbone viene poi volatilizzato, cioè distillato, nel gassificatore o anche in un recipiente separato. La maggior quantità possibile dei prodotti della distillazione deve essere trattenuta allo scopo di ridurre il costo totale della gassificazione. Poiché questi gas si sviluppano dal carbone durante la prima fase del processo e reagiscono con l'ossigeno puro o anche con quello contenuto nell'aria piú facilmente di quanto non faccia il carbone stesso, l'obiettivo di trattenere i gas di distillazione nel prodotto finale può richiedere particolari accorgimenti per assicurare che i gas non vengano in contatto con l'ossigeno e con il vapore iniettati nel sistema successivamente.

La maggior parte dei processi adatti a fabbricare un gas con un elevato potere calorifico si riferiscono sempre al vapore, all'ossigeno e al carbone come materie prime. L'ossigeno reagisce con parte del carbone per fornire il calore necessario e far salire la temperatura fino a un livello sufficiente a spostare il bilancio della reazione in modo che il carbonio reagisca rapidamente con il vapore e produca una miscela di monossido di carbonio e di ossigeno. Se si volesse far uso dell'aria al posto dell'ossigeno allo scopo di risparmiare il costo elevato dell'ossigeno, l'azoto presente nell'aria diluirebbe necessariamente il gas cosí ottenuto; pertanto alcuni processi prevedono l'impiego di un materiale intermedio per trasferire il calore necessario alla reazione. Piú precisamente, si usa un composto inorganico che viene riscaldato in un recipiente separato, con l'aria in funzione di ossidante. Questo materiale viene poi introdotto nel gassificatore, dove provoca l'aumento di temperatura richiesto perché il carbone cominci a reagire con il vapore.

Se si fabbrica idrogeno, gas di sintesi o comunque a elevato potere calorifico, il gas grezzo viene successivamente sottoposto a un trattamento che ha lo scopo di eliminare la pece e la polvere e, quindi, all'azione di un catalizzatore di scambio, che serve ad accelerare la reazione con la quale i quattro componenti del gas (anidride carbonica, idrogeno, monossido di carbonio e vapor d'acqua) raggiungono una concentrazione di equilibrio ($CO_2 + H_2 \rightleftharpoons CO + H_2O$). Si può ottenere qualunque rapporto di concentrazione si voglia fra il monossido di carbonio e l'idrogeno, semplicemente regolando la quantità di vapore e di anidride carbonica. Quando il prodotto finale desiderato è costituito dall'idrogeno, si converte tutto il monossido di carbonio in anidride carbonica per separarlo poi dal gas con uno qualunque dei processi conosciuti. Quando invece si fabbrica il gas di sintesi, che compare come materia prima per la produzione di una grande quantità di composti chimici, il rapporto fra monossido di carbonio e idrogeno viene regolato in modo da ottenere la composizione ottimale in vista del successivo processo chimico cui deve essere sottoposto.

Quando ci si prefigge come obiettivo di ottenere un gas di elevato potere calorifico, si regola il rapporto fra idrogeno e monossido di carbonio al valore piú adatto a questo scopo. Il gas viene quindi purificato per asportare i composti di zolfo che potrebbero essere presenti. Il motivo che sta alla base di questo accorgimento è che tutti i catalizzatori che si sono rivelati efficienti nella successiva e importantissima fase della metanazione sono a base di nichel e pertanto molto sensibili all'avvelenamento da parte dei composti dello zolfo.

Non si è ancora costruito alcun impianto di metanazione su scala commerciale. Un impianto pilota è stato messo in opera, tuttavia, accanto a un impianto del tipo Lurgi gestito dallo Scottish Gas Board. Questo impianto pilota, che è partito nell'autunno scorso e viene gestito da un consorzio di ditte americane, dovrebbe trattare 300 mila metri cubi al giorno di gas di Lurgi purificato per sperimentare l'efficienza di diversi catalizzatori della metanazione e per trovare il grado di pulizia cui deve soddisfare il gas per mantenere il catalizzatore attivo per un periodo di tempo ragionevolmente lungo. Un secondo grave problema tecnologico è posto dalla rapida rimozione del calore liberato dalla reazione, che risulta fortemente esotermica; se non si prevedono accorgimenti per allontanare questo calore, esso può provocare il deterioramento del catalizzatore per sinterizzazione oppure per deposizione di carbonio.

Un procedimento alternativo per la realizzazione di un gas con un alto potere calorifico prevede di gassificare una parte del carbone per ottenere idrogeno e di fare quindi reagire l'idrogeno direttamente con il carbone per produrre il metano ($C + 2H_2 \rightarrow CH_4$). In questo caso non è necessario fornire calore dall'esterno, perché la reazione risulta esotermica. I prodotti della reazione vengono quindi sottoposti al processo della metanazione, per eliminare ogni traccia residua di monossido di carbonio e per conferire al gas un potere calorifico accettabile che giustifichi la distribuzione attraverso i gasdotti. I principali vantaggi della idrogenazione diretta risiedono nel fatto che essa richiede un minor numero di materiali e che è caratterizzata da una maggiore efficienza termica che non il processo della gassificazione totale seguita dalla metanazione su catalizzatore.

La realizzazione di un gas di gassogeno è un processo molto piú semplice di quello richiesto da un gas ad alto potere calorifico. Poiché il primo richiede un minor numero di stadi di lavorazione e fa uso dell'aria invece che dell'ossigeno, il costo di produzione su scala commerciale può risultare notevolmente inferiore a quello di un gas a elevato contenuto calorico. Dopo che il carbone è stato fatto reagire con l'aria, il gas grezzo e caldissimo che ne risulta, contenente dalle 1000 alle 1450 calorie al metro cubo, viene sottoposto a un primo trattamento, che ha lo scopo di rimuovere la pece e la polvere, e quindi a un secondo trattamento per la eliminazione del solfuro di idrogeno. Il prodotto finale è un gas dallo scarso potere calorifico e con un modesto contenuto di zolfo, che può essere impiegato con successo come combustibile per caldaie. Poiché una notevole percentuale del gas naturale (il 34 per cento nel 1971) viene consumato dalle industrie, ogni sostituzione con il gas di gassogeno realizzato dal carbone potrebbe ridurre drasticamente o persino eliminare i problemi creati dall'attuale carenza di gas naturale.

I due nuovi processi che hanno raggiunto lo stadio di sviluppo dell'impianto pilota sono il processo HYGAS, che è in via di sviluppo presso l'Institute of Gas Technology, e il processo «accettore» di CO_2, che costituisce un progetto della Consolidation Coal Company. Altri processi, che sono vicini a questa fase di sviluppo, comprendono il processo BI-GAS della Bitominous Coal Research, Inc., e il processo Synthane sviluppato dall'US Bureau of Mines. Tutti i processi in fase di ricerca possono essere classificati in molti modi: secondo il metodo seguito per fornire il calore necessario alla reazione di gassificazione (riscaldamento interno o esterno); secondo il metodo usato per avere il contatto fra i reagenti (focolare fisso, fluidificato, oppure sospensione nel mezzo gassificante); secondo la direzione del flusso dei reagenti (secondo corrente oppure controcorrente); secondo la natura del mezzo gassificante (idrogeno o vapore piú ossigeno, aria oppure ossigeno arricchito); secondo le condizioni del residuo rimosso (che può essere una cenere liquida, oppure una cenere secca). Praticamente tutte le combinazioni possibili secondo queste classificazioni sono state studiate per la gassificazione del carbone.

In assenza di grandi impianti commerciali funzionanti, il costo della gassificazione del carbone può essere valutato soltanto grossolanamente, e tali stime devono essere considerate con le opportune cautele. In una presentazione al Comitato per la gassificazione del carbone della National Academy of Sciences, le rappresentanze del processo Lurgi stimarono che nelle condizioni grosso modo esistenti all'inizio del 1972 il capitale necessario per un gassificatore della portata di 7 milioni di metri cubi di gas al giorno sarebbe stato dell'ordine di una cifra compresa fra i 120 e i 130 miliardi di lire. Tali stime non comprendevano il capitale lavorativo, gli interessi passivi nel periodo necessario alla costruzione né le spese di avviamento; questi fattori avrebbero incrementato il costo portandolo a circa 170 miliardi, secondo conti basati sul valore del denaro nel 1970. Nell'autunno del 1973 la El Paso Natural Gas Company fece una stima del costo per un impianto del tutto simile, ottenendo un risultato superiore ai 270 miliardi. Tuttavia, tutti i processi elaborati piú di recente presentano dei costi di investimento stimati un poco inferiori a quelli del processo Lurgi (*si veda la figura nella pagina a fronte*).

Una tecnologia completamente diversa, che tuttavia sta richiamando l'interesse su di sé, è la gassificazione del carbone sotto terra. L'idea non è nuova, e una grande quantità di ricerche è stata effettuata intorno a essa in molti paesi fino a circa il 1960. I nuovi problemi sorti con la rarefazione delle sorgenti di energia hanno focalizzato nuovamente l'attenzione su di essa.

Qualsiasi processo di questo tipo richiede, oltre a particolari caratteristiche geologiche del giacimento, prima di tutto la formazione nel carbone di una permeabilità sufficiente a consentire che un flusso d'aria possa scorrere da un punto all'altro nella vena carbonifera senza eccessive perdite di pressione. Allora si può infiammare il carbone dando inizio al processo di gassificazione. Una parte del carbone brucia a contatto dell'aria iniettata, producendo cosí una miscela di monossido di carbonio e di anidride carbonica, oltre naturalmente al calore. Se si riesce a mantenere un contatto abbastanza buono fra il gas e il solido, l'anidride carbonica ad alta temperatura si riduce a monossido di carbonio a causa della reazione che si sviluppa con il carbone incandescente. Il prodotto che ne risulta è un gas di gassogeno.

Nel corso di vari esperimenti sulla gassificazione sotterranea del carbone si è cercato di ottenere la permeabilità necessaria attraverso un certo numero di metodi che implicano la perforazione di lunghi tunnel oppure secondo una delle varianti del metodo della filtrazione; tale metodo prevede che vengano praticati due fori dalla superficie della terra fino alla vena carbonifera e che vengano quindi connessi in modo che l'aria possa scorrere fra di essi. Questi esperimenti sono stati capaci di fornire un gas combustibile, ma non riescono a funzionare in modo sufficientemente continuo e il gas che si sviluppa non ha un potere calorifico costante. La maggior parte delle prove ha fornito come risultato uno scarso sfruttamento del carbone disponibile e non si è riusciti a raggiungere quel livello di controllo sul contatto gas-solido che risulta necessario per produrre in modo efficiente un gas dalle elevate caratteristiche. Molti altri problemi, come l'abbassamento del terreno e la contaminazione della vena da parte dell'acqua presente nel sottosuolo, non sono stati ancora risolti.

Nel frattempo sono stati proposti molti metodi nuovi, che comprendono la frammentazione del carbone per mezzo degli esplosivi e la perforazione di una serie di fori interconnessi ottenuta operando con i laser. Certo è che se si riuscirà a sviluppare un metodo efficiente, si potrà estrarre l'energia dal carbone che giace sotto terra con l'impegno di pochissima mano d'opera sotterranea o addirittura niente del tutto. Per di piú si potrebbe riuscire a estrarre l'energia contenuta nelle vene di carbone che sono troppo spesse, troppo sottili o anche di qualità troppo scadente per poter essere sfruttate secondo i metodi convenzionali.

La citocromossidasi, un enzima presente nella membrana dei mitocondri, può venire isolata per essere esaminata al microscopio elettronico e per studi di diffrazione dei raggi X. Nello stato ossidato (*in alto*) la proteina è organizzata in un lattice cristallino che si dissolve quando la proteina è allo stato ridotto (*in basso*), ma riappare quando la sostanza riducente viene eliminata. La molecola della citocromossidasi è ovale e il suo asse maggiore ha una lunghezza di circa 80 angstrom.

Un modello dinamico della membrana cellulare

Gli involucri che circondano le cellule, i nuclei e gli organelli cellulari sono un sottile insieme di molecole lipidiche e proteiche. Le loro funzioni dipendono dal modo in cui le proteine sono legate alla membrana

di Roderick A. Capaldi

L'unità fondamentale del tessuto vivente è la cellula. Negli ultimi anni è stato chiaramente dimostrato che un componente della cellula, la membrana, svolge un ruolo essenziale in quasi tutte le attività cellulari. La membrana citoplasmatica, che rappresenta il rivestimento esterno, regola l'ambiente interno della cellula e consente il trasporto di sostanze all'interno e all'esterno di essa. Le membrane interne che racchiudono il nucleo della cellula e vari organelli cellulari come microsomi, mitocondri e i cloroplasti delle piante, hanno un ruolo altrettanto importante. Per esempio, la membrana mitocondriale è il luogo dove si forma l'adenosintrifosfato (ATP); quindi questa membrana fornisce il combustibile per tutti i processi metabolici della cellula. Analogamente, la membrana dei cloroplasti è il luogo dove avviene la fotosintesi, il processo mediante il quale l'energia solare viene catturata in una forma che può essere utilizzata dalle cellule viventi. Viene spontaneo quindi chiedersi in che modo sono fatte le membrane per poter svolgere compiti così differenti.

C'è una grande quantità di informazioni per quanto riguarda la struttura fondamentale delle membrane. Un fatto emerso recentemente è che le membrane citoplasmatiche e quelle interne sono essenzialmente uguali; entrambe sono costituite da proteine e lipidi. Nelle cellule di mammiferi sono presenti anche piccole quantità di carboidrati, associati o con proteine, per formare per esempio le glicoproteine, cioè proteine portatrici di carboidrati, o con lipidi sotto forma di glicolipidi.

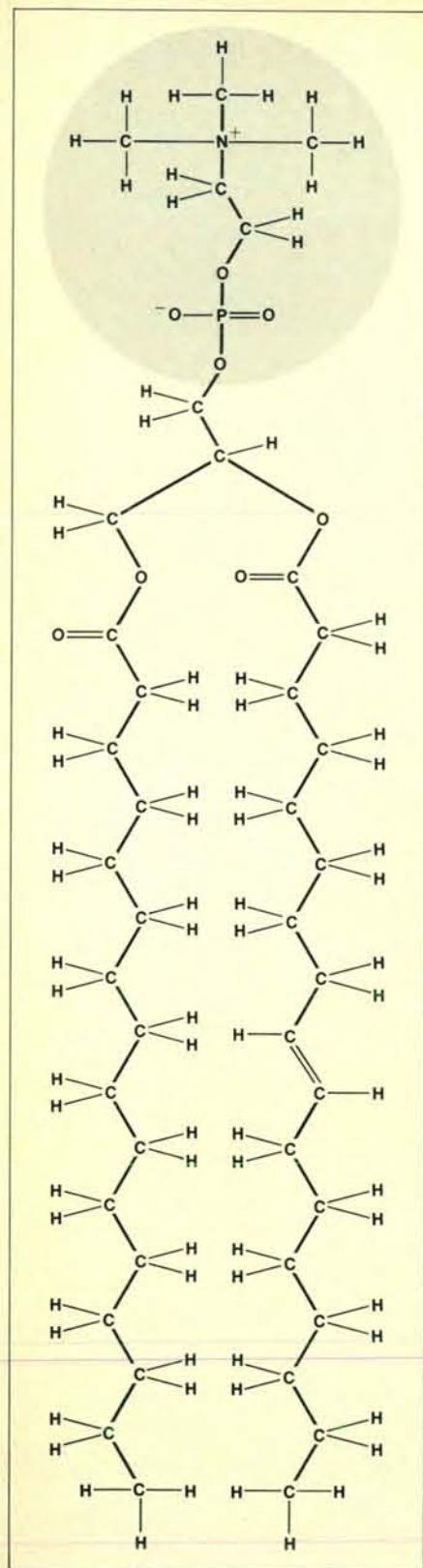
I lipidi costituiscono circa la metà della massa della maggior parte delle membrane. Nelle membrane interne, i lipidi sono costituiti quasi esclusivamente da fosfolipidi. Al contrario, le membrane citoplasmatiche, oltre ai fosfolipidi, contengono sia glicolipidi che lipidi neutri, o privi di carica. Per esempio, fino al 30 per cento dei lipidi presenti nella membrana dei globuli rossi è rappresentato da un solo tipo di lipide neutro: il colesterolo.

Le singole molecole lipidiche hanno una testa e due code (*si veda la figura nella pagina seguente*). Il punto di incontro tra la testa e le code, che C. Fred Fox definisce spina dorsale, è una molecola di glicerolo. Le code che scendono dalla spina dorsale sono delle lunghe catene di acidi grassi. La struttura di queste catene è molto simile a quella delle molecole di olio, e come l'olio e l'acqua tendono a separarsi in fasi diverse quando vengono mescolati, così le code delle molecole fosfolipidiche tendono ad allontanarsi dall'acqua. Per questo vengono definite idrofobe. D'altro canto, le teste delle molecole fosfolipidiche sono solubili in acqua e sono definite idrofile. Molecole di questo tipo, con una estremità idrofoba e una idrofila, si chiamano anfipatiche. Anche i glicolipidi e in certa misura i lipidi neutri sono anfipatici.

Nelle membrane i lipidi sono disposti in modo adeguato alle loro caratteristiche affinità. Essi formano un doppio strato, disposto in modo che le teste idrofile costituiscono le superfici superiore e inferiore della membrana e le code idrofobe rimangono all'interno della membrana (*si veda la figura in alto a pagina 19*). Il doppio strato lipidico ha uno spessore di 45 angstrom. Esso rappresenta l'intelaiatura strutturale della membrana. Costituisce anche il punto di attacco per gli altri più importanti componenti della membrana: le proteine.

Le proteine e le glicoproteine hanno ruoli molteplici nella membrana. Esse possono contribuire alla sua integrità strutturale, possono agire come enzimi o possono funzionare come « pompe » spingendo il materiale dentro e fuori dalle cellule e dagli organelli. È proprio la diversa attività delle proteine che dà a ogni particolare tipo di membrana il carattere che la contraddistingue.

Numerosi sono i metodi per determinare quali proteine siano presenti in una data membrana. Uno è quello di saggiare le varie attività enzimatiche della membrana. Un altro è quello di identificare le proteine a seconda del peso molecolare impiegando la tecnica dell'elettroforesi su gel. Le proteine sono formate da lunghe catene di amminoacidi, chiamate polipeptidi. Alcune proteine hanno soltanto una catena polipeptidica, altre hanno molte catene polipeptidiche strettamente associate le une alle altre. Nelle preparazioni per elettroforesi su gel, la proteina viene demolita nelle catene polipeptidiche che la compongono mediante esposizione a un detergente, il sodio dodecilsolfato. Le catene vengono poi trasferite in un gel di poliacrilammide su cui si crea una differenza di potenziale elettrico. Esse migrano nel gel in risposta alla differenza di potenziale a una velocità proporzionale al loro peso molecolare; più è basso il peso del polipeptide, più velocemente esso si muove. Il gel viene poi colorato con una sostanza specifica per le proteine, per esempio blu brillante di kumasi e ne viene misurato l'assorbimento. Quando con questo metodo, si identificano le proteine associate con la membrana dei globuli rossi, vengono messi in evidenza polipeptidi con un peso molecolare che varia da 255 000 a 12 500.

I due componenti polipeptidici più pesanti, con peso molecolare di 255 000 e 220 000, sono noti nel loro insieme come spettrina (Vincent T. Marchesi della Scuola di medicina dell'Univer-

Struttura di un tipico fosfolipide, la fosfatidilcolina, con una testa idrofila e una doppia coda idrofoba. Circa la metà della massa di una membrana cellulare di mammifero è costituita da varie molecole di lipidi che formano l'intelaiatura strutturale della membrana. Le loro code di acidi grassi possono essere sature (*a sinistra*) con un atomo di idrogeno legato a ogni legame carbonioso, o insature (*a destra*) con un doppio legame tra due atomi di carbonio.

sità di Yale, ha scelto questo nome perché ha isolato i polipeptidi dai cosiddetti «spettri», le membrane dei globuli rossi lisati cioè privati della emoglobina). La spettrina costituisce un terzo di tutta la parte proteica della membrana cellulare dei globuli rossi. Un altro terzo della proteina si trova in una diffusa banda di assorbimento con un peso molecolare di circa 90 000. Questa banda contiene diversi polipeptidi, compreso un costituente avente peso molecolare 87 000, sempre presente in ogni molecola di spettrina. Le proteine con peso molecolare inferiore a 70 000 costituiscono il rimanente terzo delle proteine della membrana. La membrana dei globuli rossi è un caso insolito per il gran numero di proteine ad alto peso molecolare che sono in essa contenute. Per fare un paragone, quasi tutti i polipeptidi delle membrane mitocondriali hanno un peso inferiore a 70 000.

Le proteine della membrana si possono dividere in due classi secondo la loro posizione rispetto ai lipidi della membrana. Una classe è costituita dalle molecole proteiche «estrinseche» associate cioè solo con la superficie esterna o interna della membrana. La seconda classe è costituita da proteine «intrinseche» che penetrano effettivamente all'interno della superficie della membrana. Queste proteine si insinuano nel doppio strato lipidico e a volte lo attraversano completamente (*si veda l'illustrazione in alto nella pagina a fronte*).

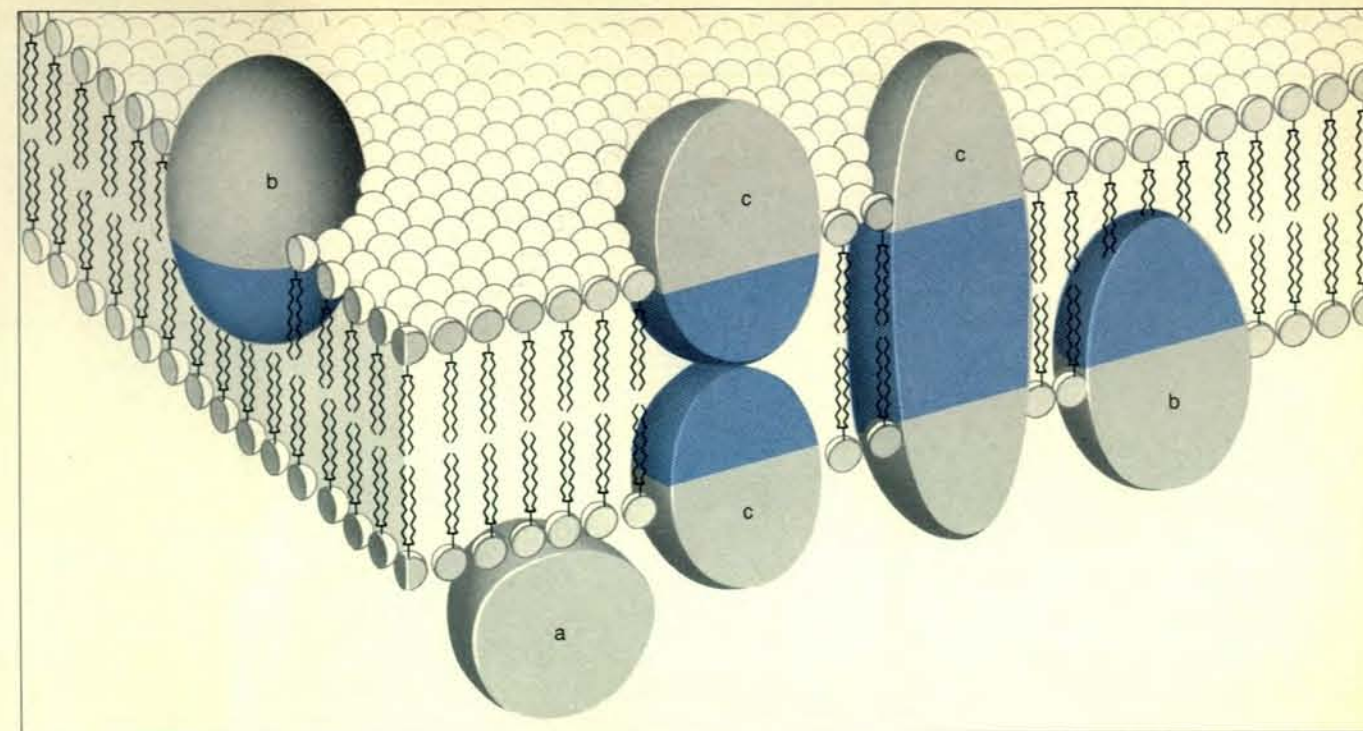
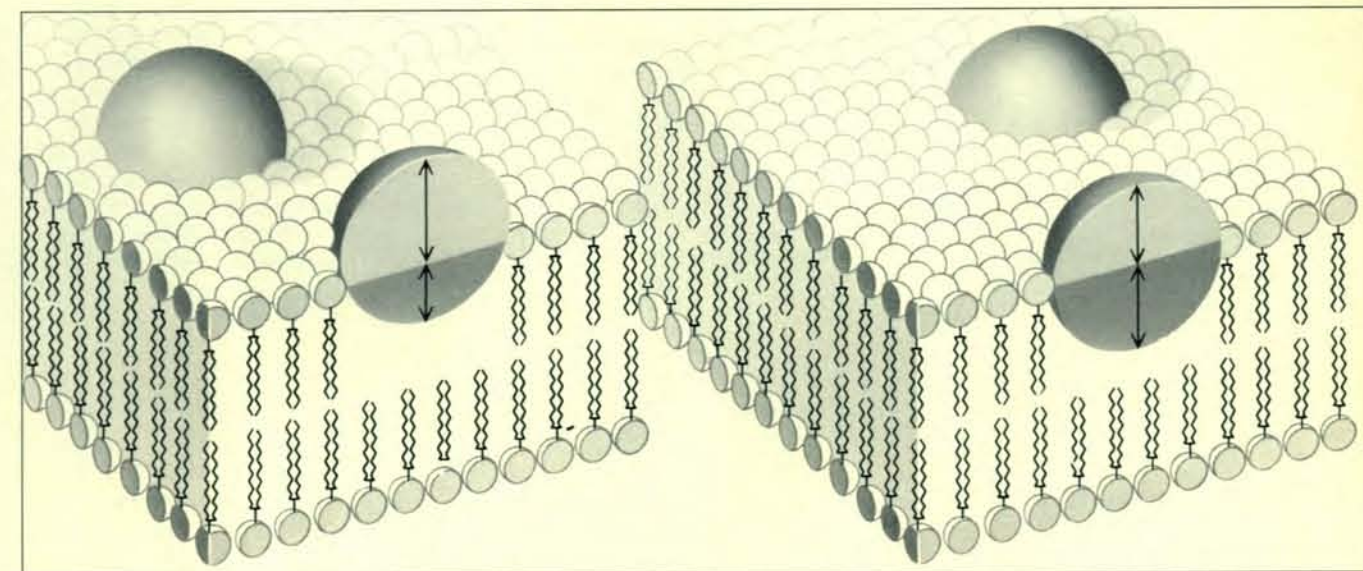
La classe di appartenenza di una proteina della membrana si determina in base alle sue proprietà chimiche o sulla base di altri tipi di analisi, come la diffrazione dei raggi X o la microscopia elettronica. Per esempio, le proteine del primo tipo si possono facilmente rimuovere dalle membrane con metodi di dissociazione chimica, mentre le proteine del secondo tipo sono parte integrante della membrana ed è molto piú difficile rimuoverle.

Due proteine estrinseche che sono visibili al microscopio elettronico sono l'enzima ATPasi, che si trova nella membrana dei mitocondri, e la spettrina, il polipeptide della membrana dei globuli rossi. Fotografie al microscopio elettronico di membrane di mitocondri, mostrano chiaramente vari corpi chiamati «teste» che emergono sulla superficie della membrana; sono molecole di ATPasi. Analogamente, in fotografie al microscopio elettronico di globuli rossi svuotati dall'emoglobina, la «peluria» che circonda l'interno della membrana è composta da polipeptidi di spettrina.

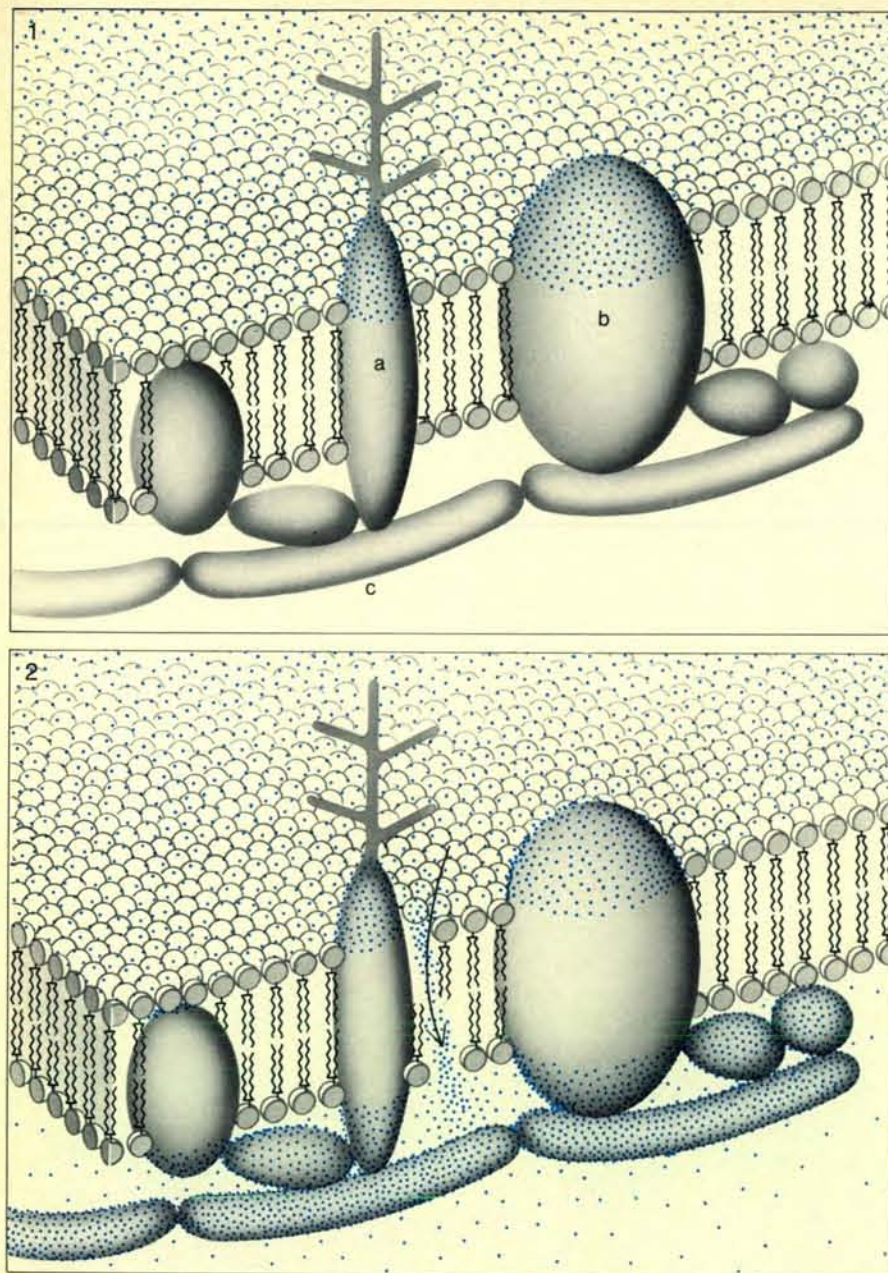
Una proteina intrinseca che è stata ampiamente studiata è la rodopsina, l'unica proteina presente nelle membrane dei dischi che occupano i segmenti esterni dei bastoncelli della retina. J. Kent Blasie e i suoi collaboratori della Scuola di medicina dell'Università di Pennsylvania, lavorando con tecniche di diffrazione dei raggi X, hanno trovato che la molecola della rodopsina è globulare e ha un diametro di circa 42 angstrom. Quando i bastoncelli della retina sono al buio, le molecole della rodopsina della membrana del disco, sono immerse per circa un terzo del loro diametro nella superficie esterna della membrana. Quando i bastoncelli sono illuminati, le molecole di rodopsina si insinuano piú profondamente nella membrana fino a essere immerse a metà.

Insieme ad alcuni colleghi del laboratorio di David E. Green all'Istituto di ricerche sugli enzimi dell'Università del Wisconsin, ho esaminato la organizzazione di un'altra proteina intrinseca, la citocromossidasi, un enzima della membrana mitocondriale che è l'elemento terminale della catena di trasferimento degli elettroni implicata nella sintesi dell'ATP. Un ostacolo nelle ricerche sulle membrane proteiche, è il fatto che la maggior parte delle membrane contiene una miscela eterogenea di proteine che comprende sia proteine estrinseche che intrinseche. La maggior parte dei metodi che possono essere impiegati per esaminare la struttura delle membrane, come la diffrazione dei raggi X, sono metodi con risultati intermedi; i dati che forniscono rappresentano soltanto le proprietà medie di tutte le proteine del campione, mentre in realtà vogliamo conoscere le caratteristiche delle singole proteine della membrana. Questa è la ragione per cui la membrana dei dischi della retina, con la sua unica proteina, la rodopsina, è un soggetto molto diffuso in questo tipo di ricerche.

Fortunatamente per noi è possibile separare le citocromossidasi dalle altre proteine della membrana dei mitocondri. Quando l'enzima separato viene posto in sospensione con i lipidi, queste due sostanze interagiscono e si raccolgono in vescicole a forma di sacco che sono in realtà vere membrane artificiali. Le molecole della citocromossidasi nelle vescicole artificiali hanno le stesse proprietà enzimatiche che esse presentano nella membrana mitocondriale normale e le vescicole hanno la stessa struttura delle normali membrane. Il vantaggio, in questo caso, è che un insieme eterogeneo di proteine è stato ridotto a una sola proteina.

Abbiamo usato queste vescicole di membrana, come modello per lo stu-

L'intelaiatura delle membrane cellulari è costituita da un doppio strato di lipidi le cui teste idrofile formano le superfici interne ed esterne e le code idrofobe si uniscono al centro della membrana; il doppio strato ha uno spessore di circa 45 angstrom. Le proteine, gli altri componenti della membrana, sono di due tipi. Alcune (*a*) si trovano su entrambe le superfici della membrana. Le altre penetrano nella membrana per un breve tratto (*b*) o la attraversano completamente da sole o a coppie (*c*).

Il grado di penetrazione può variare a seconda delle circostanze. J. Kent Blasie e i suoi collaboratori hanno dimostrato che la rodopsina, la sola proteina della membrana dei dischi dei bastoncelli della retina, è una molecola globulare con un diametro di 42 angstrom. Al buio, un terzo della molecola è immerso nei lipidi della membrana (*a sinistra*): alla luce, metà.

dio della proteina isolata nelle sue relazioni geometriche con il doppio strato lipidico. Come sostanza in grado di trasportare elettroni, la citocromossidasi esiste sia nello stato ossidato che in quello ridotto. Allo stato ossidato (e a un determinato rapporto lipidi-proteine) l'enzima è organizzato in un lattice cristallino visibile al microscopio elettronico e che può essere analizzato con diffrazione dei raggi X (*si veda l'illustrazione a pagina 16*). Utilizzando èntrambi i tipi di dati, abbiamo trovato che le singole molecole di citocromossidasi sono lunghe circa 55 angstrom, larghe 60 angstrom e profonde da 80 a 85 angstrom. La profondità è sufficiente per permettere alla molecola di penetrare completamente nel doppio strato che ha uno spessore di 45 angstrom, lasciando una estremità che sporge dalla superficie esterna della membrana e l'altra che sporge in modo simile nella superficie interna. Si tratta quindi di una proteina che attraversa la membrana estendendosi attraverso il doppio strato.

Le proteine intrinseche hanno un ambiente molto diverso da quello delle proteine estrinseche che sono essen-

Marcando le proteine è possibile mettere in evidenza la distribuzione asimmetrica delle molecole proteiche associate con la membrana dei globuli rossi. Quando sulla superficie della cellula (*1*) si applica un reattivo chimico che non può passare attraverso il doppio strato lipidico, vengono marcate soltanto due proteine le cui estremità si estendono al di sopra della superficie (*a*, *b*). Quando si aggiunge una sostanza che provoca lisi, rendendo la membrana permeabile (*2*), vengono marcate tutte le molecole dell'altro lato della membrana, compresa la spettrina (*c*) che trattiene le altre al loro posto.

zialmente circondate da acqua. Soltanto una parte di una proteina intrinseca è esposta all'acqua; il resto è immerso in un liquido oleoso. Allo scopo di mantenere la stabilità in questo insolito duplice ambiente, la proteina deve essere anfipatica, come le molecole dei fosfolipidi. Le parti della molecola esposte all'acqua devono possedere la maggior parte dei sette amminoacidi idrofili: lisina, istidina, arginina, acido aspartico, acido glutammico, serina e treonina. Le parti della molecola che si addentrano nel doppio strato, invece, devono essere per lo piú costituite da amminoacidi idrofobi.

Quanto piú in profondità viene a trovarsi la proteina nel doppio strato, tanto minore sarà la superficie circondata dall'acqua in grado di accogliere gli amminoacidi idrofili nella molecola. Non sorprende perciò che molte proteine intrinseche contengano relativamente pochi amminoacidi idrofili, pur mostrando la tendenza a comportarsi come idrofobe. Per esempio, Alexander Tzagaloff dell'Istituto di ricerca sulla salute pubblica di New York, ha isolato una proteina intrinseca delle membrane interne dei mitocondri con un peso molecolare di 10 000. Il contenuto in amminoacidi è per il 20 per cento idrofilo e per l'80 per cento idrofobo. Questo fatto è in netto contrasto con le proteine citoplasmatiche ed estrinseche che hanno in media il 47 per cento di amminoacidi idrofili e il 53 per cento di amminoacidi idrofobi. Un'altra proteina intrinseca molto idrofoba è la proteina di Folch-Lees, che può essere isolata dalla mielina che riveste il nervo sciatico; i suoi amminoacidi sono idrofili per il 29 per cento e idrofobi per il 71 per cento. Le proteine intrinseche, rodopsina e citocromossidasi, sono entrambe idrofobe; gli amminoacidi della rodopsina sono idrofili per il 36 per cento e idrofobi per il 64 per cento e quelli della citocromossidasi sono idrofili per il 37 per cento e idrofobi per il 63 per cento.

Recentemente sono state studiate due interessanti proteine intrinseche, le quali, mentre nella composizione generale non sono eccessivamente idrofobe hanno zone a catena polipeptidica molto ricche di amminoacidi idrofobi. Una è il citocromo b_5, una proteina isolata dalla membrana microsomale di cellule epatiche. Phillip Strittmatter e Lawrence Spatz dell'Università del Connecticut hanno dimostrato che questa proteina, formata da un singolo polipeptide, è ripiegata a una estremità in una parte globulare che sporge sulla superficie della membrana ed è in gran parte ricoperta da amminoacidi idrofili. La catena polipeptidica continua al di là della porzione globulare con una « coda » di circa 60 amminoacidi, quasi tutti idrofobi. La coda penetra nel doppio strato e serve per ancorare alla membrana la parte globulare ed enzimaticamente attiva della molecola.

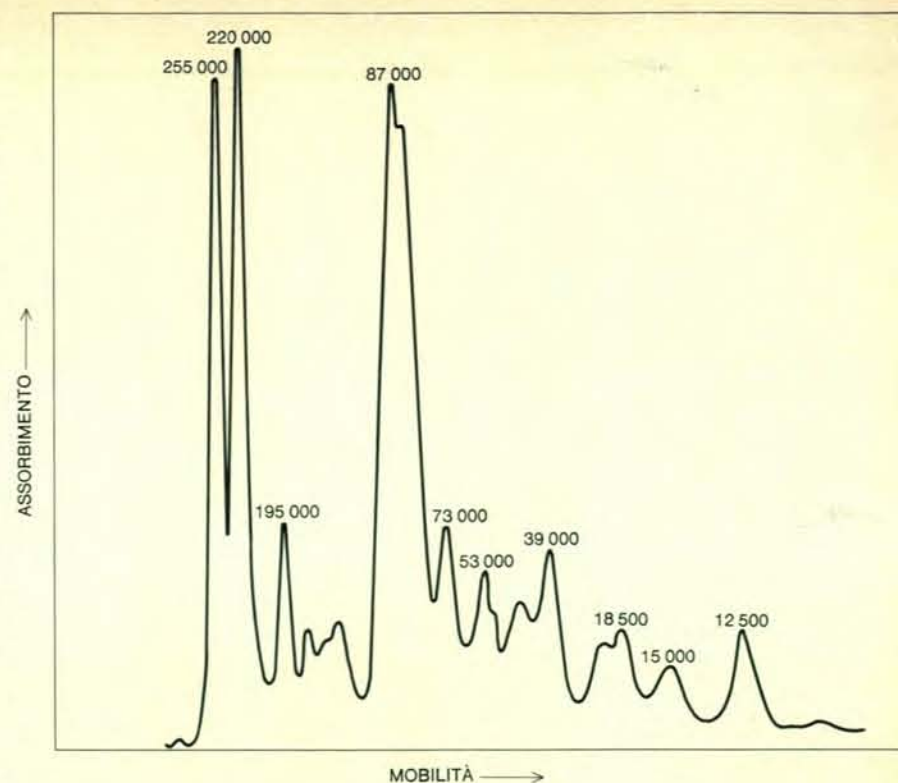
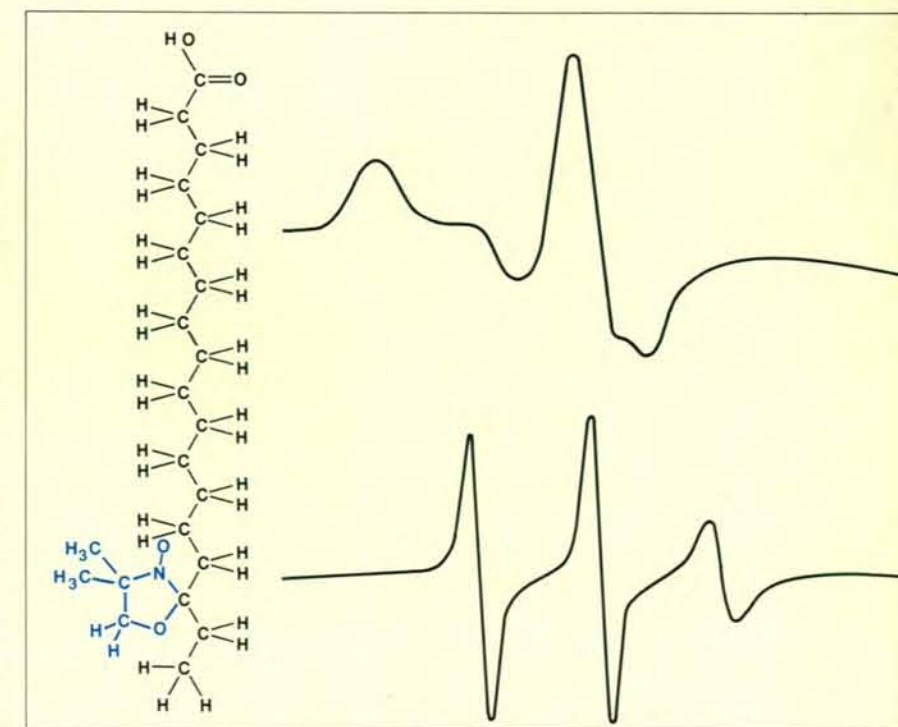
La seconda proteina intrinseca che ha una struttura piuttosto insolita, è la piú importante glicoproteina trovata nella membrana dei globuli rossi. Questa molecola è stata studiata da Marchesi e da vari suoi colleghi, prima al National Institutes of Health e piú recentemente a Yale. Anche questa è formata da un'unica catena polipeptidica. Una estremità della catena, che contiene tutti i carboidrati associati alla molecola, è formata per lo piú da amminoacidi idrofili; questa estremità è esposta all'acqua e si trova sulla superficie esterna della membrana. L'altra estremità della catena, che contiene anch'essa amminoacidi idrofili, si estende nell'interno acquoso del globulo rosso. La parte intermedia della catena è formata da circa 30 amminoacidi. Essi sono quasi esclusivamente idrofobi e si trovano all'interno del doppio strato lipidico della membrana.

Poiché tutti i carboidrati della molecola glicoproteica sono disposti sulla superficie esterna della membrana del globulo rosso, la membrana è asimmetrica. Inoltre, studi con metodi di marcatura, nei quali si usano reattivi chimici in grado di marcare le proteine disponibili mostrano che l'asimmetria della membrana del globulo rosso va oltre il limite del contenuto in carboidrati. I reattivi chimici che non possono penetrare nel doppio strato lipidico della membrana riescono a marcare due delle proteine presenti nella membrana del globulo rosso intatto. Una è la glicoproteina, l'altra è una proteina a peso molecolare 87 000; solo queste due proteine sono esposte alla superficie esterna della membrana. Tuttavia, quando il globulo rosso viene sottoposto a lisi, e quindi reso permeabile a marcatori chimici, tutte le proteine della membrana vengono marcate, e ciò sta a indicare che la maggior parte delle proteine della membrana del globulo rosso è localizzata nella superficie interna della membrana.

La membrana del globulo rosso non è la sola ad avere un'organizzazione asimmetrica. Le tecniche di marcatura hanno dimostrato che la membrana interna dei mitocondri è organizzata in modo simile. Le molecole della proteina, note come teste, in effetti ATPasi, si trovano esclusivamente nella matrice o parte interna della membrana e le molecole del citocromo *c* si trovano soltanto sulla superficie esterna. Come si può prevedere, le molecole della citocromossidasi che sono esposte sia alla superficie intracristale, sia a quella della matrice della membrana, si possono marcare su entrambi i lati.

Per riassumere il quadro fin qui tracciato della membrana, osserviamo che i lipidi che costituiscono circa la metà della massa di una membrana, sono organizzati in un sottile doppio strato e che le proteine della membrana si trovano o sopra o vicine ai due lati del doppio strato, oppure completamente lo attraversano.

Il quadro è accurato, ma non completo. Uno dei piú importanti progressi nello studio della cellula, compiuto negli ultimi anni, è l'aver capito che le membrane non sono affatto statiche, sia i lipidi che le proteine hanno notevole libertà di movimento.

Considerando, in primo luogo, la mobilità dei lipidi, si vedrà che un doppio strato lipidico è fluido o rigido in dipendenza di due fattori: primo, il grado di saturazione delle code lipidiche,

L'elettroforesi su gel permette di determinare l'identità delle proteine nella membrana dei globuli rossi. Nell'esempio qui riportato, le proteine sono state marcate e esaminate usando un densitometro, dopo che quelle con peso molecolare piú basso erano migrate piú velocemente attraverso il gel in risposta a una differenza di potenziale elettrico. Le due proteine con il peso molecolare piú elevato (*a sinistra*) formano il dimero chiamato nel suo complesso spettrina. L'esame mostra almeno altre 10 bande di assorbimento corrispondenti ad altre proteine con un peso molecolare piú basso.

La mobilità dei lipidi del doppio strato della membrana può essere determinata con spettroscopia ESR. Un gruppo di riferimento sensibile al movimento (*a colori*) è attaccato a un atomo di carbonio nella coda di una molecola campione (*a sinistra*). Quando la molecola viene inserita nel doppio strato, lo spettro del gruppo di riferimento varierà in rapporto al grado di mobilità del doppio strato. Se la bassa temperatura ha congelato il doppio strato in modo da immobilizzare i lipidi, si ottiene uno spettro caratteristico (*in alto a destra*). Se il doppio strato è mobile, il gruppo di riferimento nella coda del lipide dà luogo a uno spettro nettamente diverso (*in basso a destra*).

(cioè in quale misura tutti i legami carboniosi disponibili portano atomi di idrogeno) e secondo, la temperatura dell'ambiente. Ora, una parte notevole dei lipidi nelle membrane cellulari di mammiferi è insatura, cosicché la temperatura di fusione del doppio strato è al di sotto della temperatura corporea normale dei mammiferi. Quindi il doppio strato è fluido e le code degli acidi grassi delle molecole lipidiche sono libere di muoversi. La libertà di movimento è stata studiata abbastanza dettagliatamente da Harden M. Mc Connel e collaboratori all'Università di Stanford e da O. Hayes Griffith e collaboratori all'Università dell'Oregon. Entrambi i gruppi hanno usato la stessa tecnica di analisi: spettroscopia a risonanza elettronica di spin (ESR). Questo metodo implica l'attacco di un gruppo di riferimento, di solito un gruppo nitroso che ha un elettrone non appaiato, a uno degli atomi di carbonio della coda di un acido grasso di una molecola campione.

La molecola campione è o un acido stearico o un fosfolipide e il gruppo di riferimento è di solito attaccato al 5°, al 12° o al 16° atomo di carbonio lungo la coda. Nella membrana del doppio strato da esaminare vengono quindi inserite numerose molecole campione marcate; esse si adattano per lo più allo stesso modo di quanto avviene per le molecole lipidiche normali. Tuttavia, il gruppo marcato è molto sensibile al movimento e il tracciato spettroscopico registra senza errore qualsiasi movimento della parte della coda dell'acido grasso dove è attaccato il gruppo di riferimento. Fissando il gruppo di riferimento in diverse posizioni lungo la coda, si può determinare in particolare quale grado di mobilità ci sia e in quale posizione del doppio strato.

Con questo metodo i gruppi di Mc Connel e Griffith hanno dimostrato che nel doppio strato esiste un gradiente di flessibilità. La parte della coda che è più vicina alla testa della molecola lipidica (e quindi più vicina alla superficie del doppio strato) è la meno flessibile. Al contrario, la maggior flessibilità si riscontra nelle estremità delle code dei lipidi cioè in quelle porzioni della membrana più vicine al centro del doppio strato.

Nel doppio strato non tutte le molecole lipidiche sono mobili. Per esempio, le molecole proteiche penetranti influenzano la mobilità delle molecole lipidiche a esse adiacenti. In collaborazione con Griffith e P.C. Jost, abbiamo usato acidi stearici *spin labeled* (con un gruppo funzionale con un elettrone spaiato) per saggiare l'ambiente lipidico adiacente a molecole di citocromossidasi nel nostro modello di membrana.

Abbiamo trovato che le molecole di citocromossidasi, in effetti, immobilizzano una quantità sufficiente di lipidi per ricoprirsi con un singolo strato di molecole lipidiche. A questo strato intimamente associato abbiamo dato il nome di «lipide di delimitazione». È interessante notare che la quantità di questo lipide corrisponde alla quantità di lipide necessaria perché l'attività di questo enzima si svolga in modo completo. La quantità esatta del contenuto totale di lipidi che viene immobilizzata per costituire il lipide di delimitazione, dipende dal numero di molecole proteiche penetranti che una particolare membrana contiene. Nella membrana mitocondriale, che contiene molte molecole penetranti, forse quasi il 30 per cento dei lipidi della membrana viene impiegato a questo scopo.

Poiché una notevole percentuale del doppio strato della membrana è fluida, la membrana nel suo insieme ha la consistenza di un olio leggero. Non sorprende quindi che, a meno che le circostanze non lo impediscano, le molecole lipidiche e proteiche siano libere di muoversi all'interno della struttura a foglietto. Tale movimento si svolge più in direzione orizzontale che verticale; a riprova di questo è la stabilità delle configurazioni asimmetriche della membrana. In effetti, sarebbe difficile, da un punto di vista energetico, per una molecola di proteina intrinseca forzare la sua estremità altamente idrofila per penetrare attraverso la parte interna idrofoba del doppio strato, allo scopo di raggiungere l'altro lato della membrana. Lo spreco di energia sarebbe minore nel caso di una molecola lipidica, ma anche in questo caso, studi con fosfolipidi *spin labeled* indicano che lo spostamento del lipide da una parte all'altra del doppio strato avviene raramente.

Il movimento laterale delle molecole nel doppio strato è stato dimostrato per la prima volta in esperimenti condotti da David Frye e Michael Edidin dell'Università Johns Hopkins nel 1970. Essi stavano studiando le proprietà delle membrane in «supercellule», ottenute per fusione indotta di varie cellule sotto l'influenza del virus Sindai, un noto agente di fusione.

Frye e Edidin provocavano la fusione di cellule umane e di topo in colture e poi studiavano la distribuzione di alcune proteine intrinseche della membrana: i componenti antigenici dei due tipi di cellule. Il luogo dove ogni proteina era presente, poteva essere determinato marcando gli anticorpi con sostanze fluorescenti; gli anticorpi marcati in modo diverso venivano diretti o verso l'antigene umano o verso quello del topo. I ricercatori trovarono che poco dopo la fusione di due cellule, i componenti antigenici del topo e dell'uomo venivano chiaramente segregati in due zone separate della membrana della supercellula. Tuttavia, dopo che la supercellula era stata incubata per 40 minuti a 37° C, le proteine dell'uomo e del topo si erano in sostanza rimescolate tra di loro. Poiché il mescolamento era avvenuto in assenza di ATP, non si trattava evidentemente di un processo che richiedeva energia. Frye e Edidin poterono concludere soltanto che il mescolamento era un prodotto della diffusione laterale attraverso il doppio strato. Trovarono che quando la coltura veniva tenuta a 1,1°C, una temperatura sufficiente a «congelare» la membrana, le due proteine non si mescolavano.

Negli anni successivi a questo esperimento classico, si è trovato che molti altri antigeni e molecole proteiche si muovono nel doppio strato della membrana. Anche le molecole lipidiche sono altrettanto mobili come è stato dimostrato dagli studi con spettroscopia ESR condotti da McConnel e Philippe Devaux. Essi inserivano delle porzioni di fosfolipide *spin labeled* nelle membrane e registravano il tempo necessario per disperdere le molecole marcate mediante diffusione. Trovarono che i lipidi si muovevano con una velocità superiore rispetto alle molecole proteiche; essi infatti sono più piccoli delle molecole proteiche.

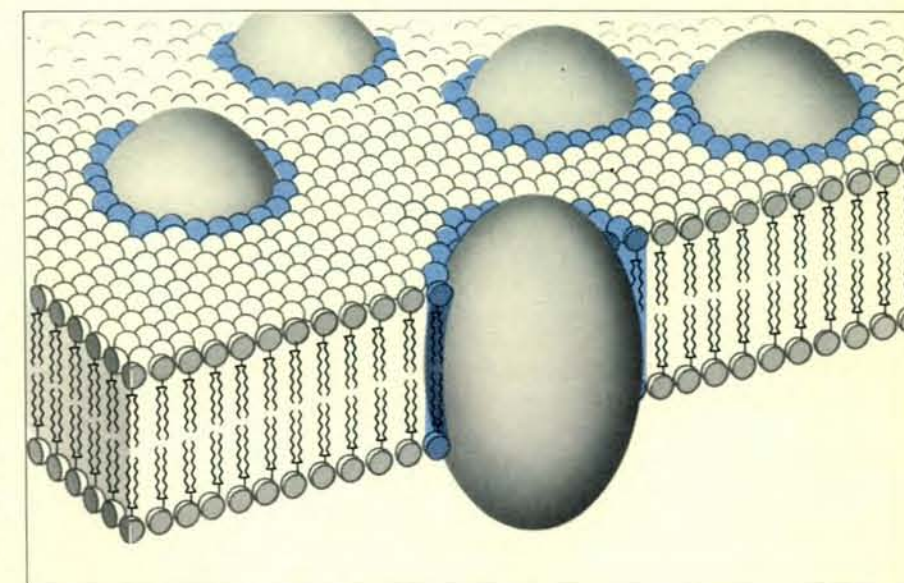
Quindi, se lipidi e proteine fossero tutti liberi di muoversi nel doppio strato della membrana, ci si dovrebbe aspettare che i due componenti fossero distribuiti a caso nella membrana. Per molte membrane, tuttavia, questo non accade. Per esempio, nella membrana delle cellule che tapezzano l'intestino, le glicoproteine sono concentrate sul lato superficiale della cellula e le proteine che pompano il sodio sono concentrate all'estremità opposta. Un altro esempio è dato dalla cellula nervosa; in questo caso, una proteina chiave, l'enzima acetilcolinesterasi, è localizzato esclusivamente nella membrana a una estremità della cellula. Alcune membrane presentano anche lattici cristallini del tipo trovato nel modello di membrana della citocromossidasi: la membrana dei cromatofori del batterio *Halobacterium halobium* ne è un esempio. Questa «membrana purpurea» contiene soltanto una proteina con una ben definita struttura a lattice. I ponti cellulari costituiscono un altro esempio. Daniel A. Goodenough della Scuola di medicina di Harvard, ha isolato queste strutture da cellule di fegato di topo. Si pensa che i ponti facilitino il trasferimento di sostanze da una cellula all'altra. Esse contengono prevalentemente una proteina e un lipide, probabilmente derivati da entrambe le cellule, organizzati in un lattice esagonale chiaramente visibile al microscopio elettronico.

Quali fattori influenzano la mobilità e quindi la distribuzione delle proteine nelle membrane? I dati finora disponibili suggeriscono che il controllo sia esercitato in modi diversi. Per esempio, sembra probabile che nella membrana interna dei mitocondri ciò che mantiene la distribuzione spaziale sia un'intima associazione tra le varie proteine intrinseche della membrana. Nei mitocondri quasi tutte le proteine sono associate con uno o con l'altro dei cinque complessi: quattro implicati nel trasporto di elettroni, e uno che sintetizza ATP. Sembra probabile che quantità proporzionali di ciascuno dei cinque complessi si combinino per formare aggregati sopramolecolari che sono ordinatamente disposti attraverso la membrana. Tali aggregati favorirebbero l'efficienza del trasporto di elettroni e l'accoppiamento del trasporto degli elettroni alla sintesi dell'ATP.

La distribuzione delle proteine nella membrana dei globuli rossi è probabilmente mantenuta in modo diverso. In questo caso, la distribuzione è controllata in modo evidente da un'interazione tra le proteine intrinseche, da un lato, e le molecole di spettrina con le proteine estrinseche che si trovano sulla superficie interna della membrana, dall'altro. Garth Nicolson dell'Istituto Salk per ricerche biologiche ha fornito un'interessante dimostrazione del ruolo svolto dalla spettrina sul controllo di tale distribuzione. Nicolson ha preparato un anticorpo per la spettrina che applicato ai globuli rossi, lisati, fa sì che le molecole di spettrina si aggreghino in modo anormale. Quando questo accade, la distribuzione delle proteine intrinseche della membrana, e in modo particolare la distribuzione della glicoproteina viene cambiata. La parte idrocarboniosa di questa molecola contiene molti gruppi acidi, chiamati acidi sialici, che legano un materiale difficilmente attraversabile dagli elettroni e facilmente visibile al microscopio elettronico. Il materiale può essere quindi usato per «colorare» la glicoproteina. Quando a una preparazione di globuli rossi si aggiunge l'anticorpo per la spettrina, la posizione delle molecole colorate è radicalmente alterata. Evidentemente, la glicoproteina (o più probabilmente un aggregato sopramolecolare in cui essa è compresa) è così strettamente legata alla spettrina sulla superficie esterna della membrana che quando si muove una, l'altra la segue.

Questi sono soltanto due dei numerosi sistemi che controllano la mobilità e la distribuzione delle proteine nelle membrane. Per esempio, un sistema completamente diverso entra in gioco per quanto riguarda la rodopsina nella membrana dei bastoncelli della retina e sembra che la mobilità e la distribuzione delle proteine di quelle membrane siano controllate anche da altri sistemi che implicano reti di microtubuli e microfilamenti situati sotto la superficie interna della membrana citoplasmatica.

Per riassumere, sembra dimostrato che in alcune membrane la distribuzione spaziale delle proteine sia rigida e ben determinata, mentre in altre membrane tale distribuzione possa variare notevolmente sotto il controllo che avviene all'interno della cellula.

Gli aggregati sopramolecolari della membrana dei globuli rossi, comprendono due proteine che attraversano la membrana. Una è una glicoproteina (*a*), l'altra è una proteina di peso molecolare 87 000 (*b*). La terza proteina dell'ipotetico aggregato è la spettrina (*c*). Le tre proteine sono legate in modo che, se una si muove, le altre la seguono.

Il lipide di delimitazione è un «involucro monomolecolare» di molecole lipidiche immobilizzate (*in colore*) che circonda le grosse molecole proteiche. La sua esistenza è stata scoperta nel corso di studi di spettroscopia ESR su membrane «modello» contenenti una sola proteina, la citocromossidasi. In membrane con più proteine penetranti, come nella membrana mitocondriale, è immobilizzato più del 30 per cento dei lipidi.

Le basi nervose del comportamento guidato dalla vista

Mediante l'applicazione di tecniche etologiche e neurofisiologiche si può dimostrare in qual modo un animale localizza uno stimolo ottico, ne identifica il significato e infine emette l'appropriata risposta motoria

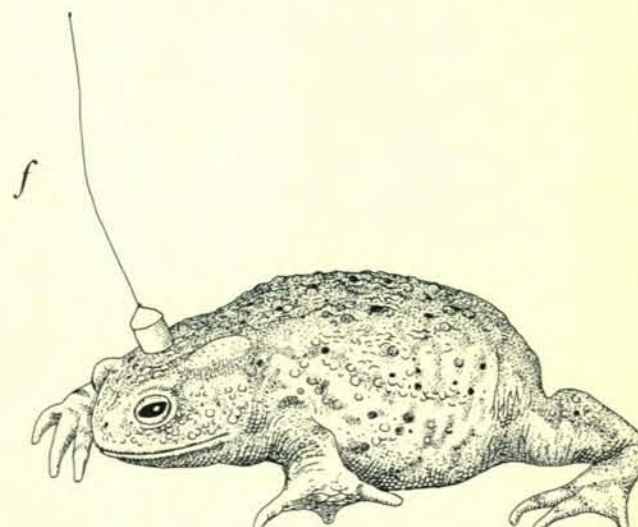
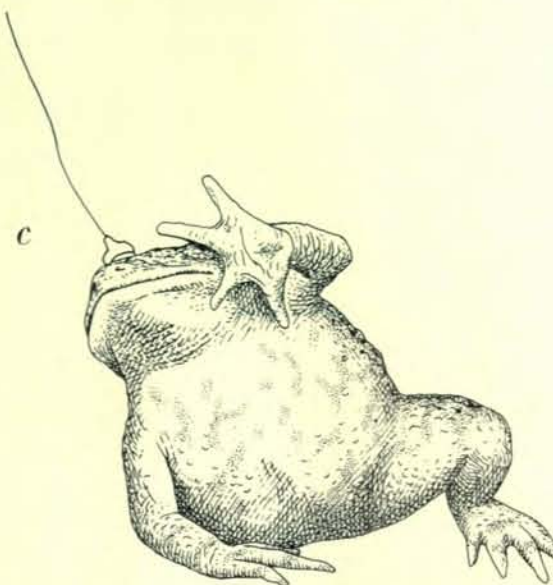
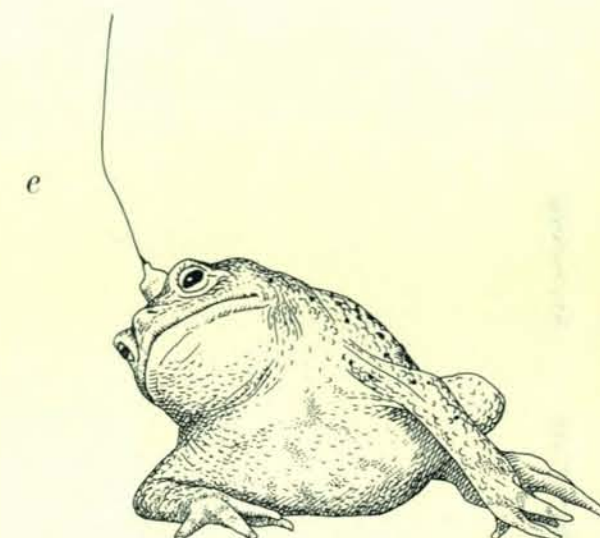
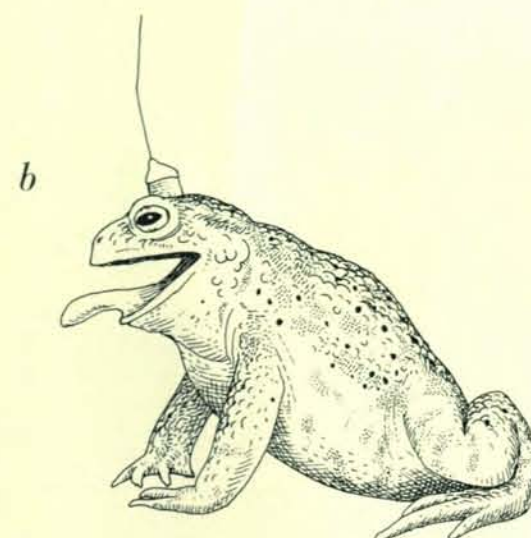
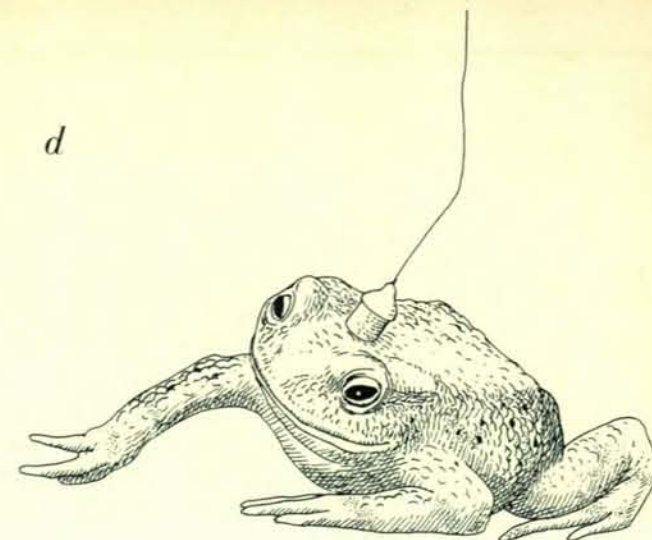
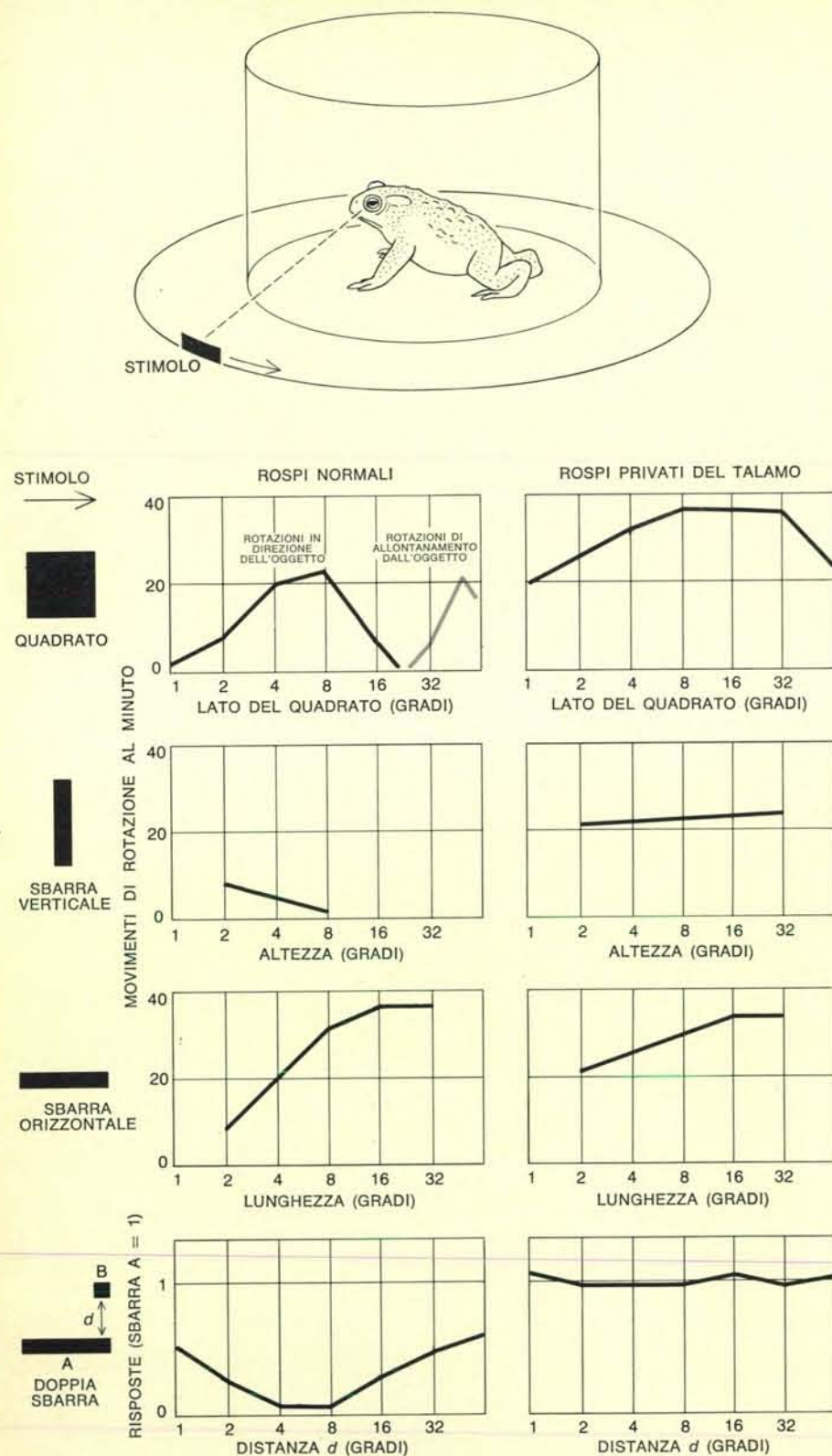
di Jörg-Peter Ewert

Gli animali vedono gli oggetti e poi agiscono sulla base di ciò che hanno visto. Quale catena di eventi connette alcuni stimoli-chiave a uno specifico schema fisso di risposte? Negli ultimi anni studiosi di diversi laboratori hanno cercato di analizzare, con una grande varietà di metodi, i meccanismi nervosi che permettono agli animali di interpretare i segnali sensoriali e selezionare la risposta più appropriata. Il modo più efficace per comprendere le basi nervose del comportamento sembra consistere nell'applicare una vasta gamma di metodi sperimentali: di associare, cioè, gli studi etologici del comportamento animale a ricerche sull'anatomia del cervello e a tecniche di stimolazione delle cellule cerebrali e di registrazione dell'attività delle singole cellule nervose.

Negli ultimi sei anni, nel mio laboratorio, prima presso la Technical University e poi presso l'Università di Kassel, abbiamo affrontato su queste basi lo studio di due comportamenti guidati dalla vista nel rospo: il comportamento di orientamento (cattura della preda) e il comportamento di evitamento (fuga). Esistono diverse valide ragioni per lavorare con il rospo. Gli anfibi sono vertebrati, e dunque ciò che noi impariamo al loro livello di integrazione comportamentale relativamente basso ci aiuta a comprendere il comportamento di vertebrati più complessi. In particolare i rospi hanno un repertorio di comportamenti limitato e facilmente esaminabile. È possibile provocare ripetutamente reazioni predeterminate in risposta a specifici stimoli quali lo schiocco con la lingua alla vista della preda, la fuga da un nemico, l'atto di allacciare la compagna durante il corteggiamento e particolari reazioni di strofinamento in risposta alla stimolazione tattile. (La variante europea della rana, al contrario, va incontro a rapide modificazioni del livello di motivazione e non è adatta a esperimenti sul comportamento.) Il rospo, infine, non si condiziona facilmente, per cui le funzioni comportamentali innate possono essere misurate in successivi esperimenti per un dato periodo di tempo, senza che esse vengano alterate in misura significativa dall'esperienza accumulata.

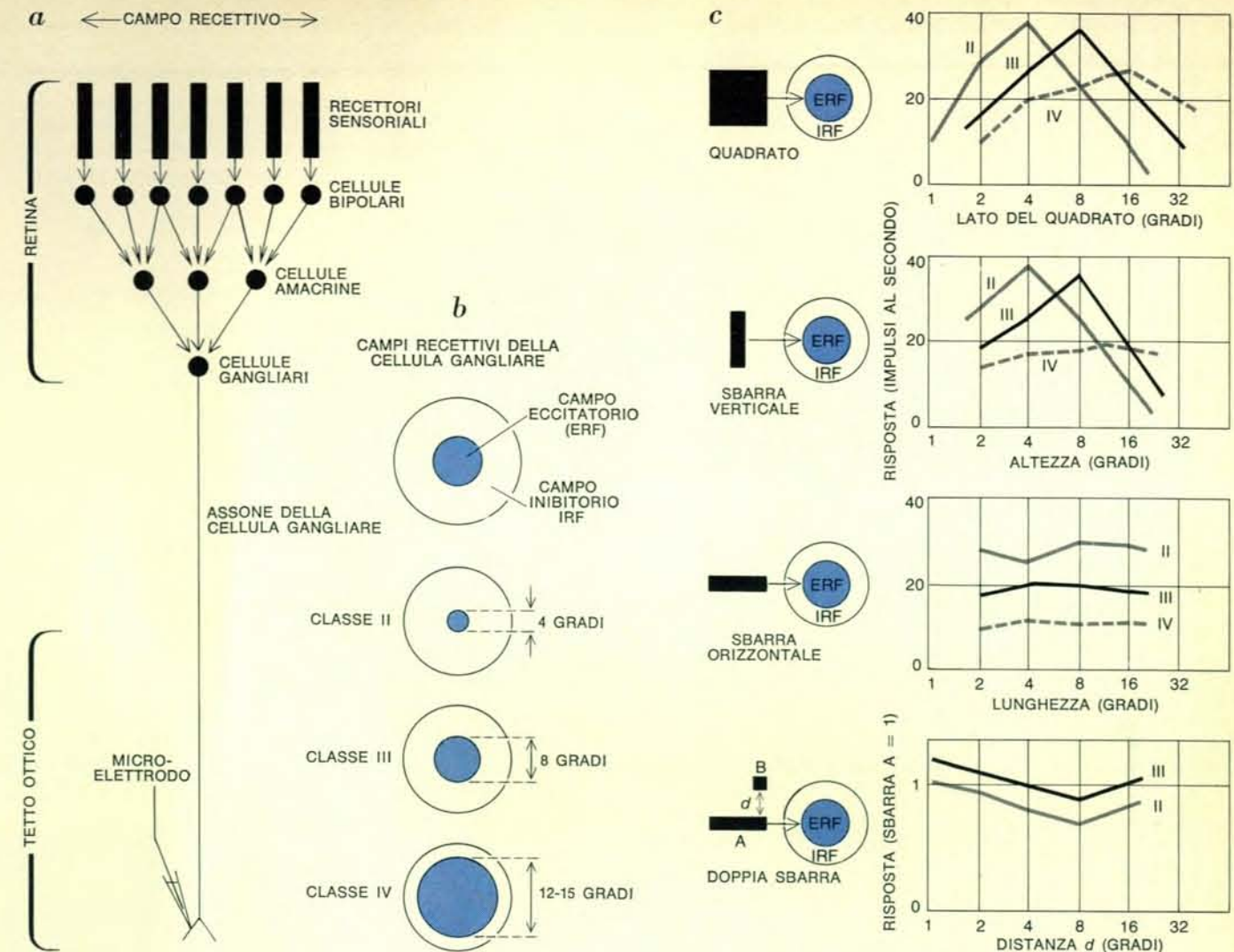
I rospi rispondono a oggetti piccoli, come per esempio un pezzo di cartone bianco fatto muovere su uno sfondo nero, con una serie di reazioni di cattura. In un primo momento si ha un orientamento verso la preda, poi una fissazione binoculare, schiocchi con la lingua, deglutizioni e pulitura della bocca. Perché la reazione di orientamento sia completa, sono necessari due processi fondamentali: l'identificazione dello stimolo e la sua localizzazione nello spazio. Il processo di identificazione determina il tipo di comportamento. Esso è in rapporto a caratteri specifici dello stimolo, quali la sua grandezza angolare, l'orientamento dei confini tra luce e buio, la velocità angolare, il grado di contrasto con lo sfondo e così via. A questo punto interviene un processo di localizzazione dello stimolo e la somma dei due processi determina la risposta motoria che può essere rappresentata o da un movimento in direzione dello stimolo, se questo viene identificato come « preda », o da un movimento di allontanamento da esso, se è stato identificato come « nemico ». Nelle pagine che seguono cercherò di analizzare le basi neurofisiologiche della identificazione del segnale, della sua localizzazione e dello scatenamento delle azioni istintive coordinate.

Il primo passo è rappresentato dall'analisi quantitativa degli stimoli-chiave per il comportamento di orientamento e di evitamento. Ciò si ottiene variando secondo un ordine preciso alcune caratteristiche dello stimolo visuale. Il rospo viene posto all'interno di un contenitore cilindrico di vetro e gli viene mostrato un piccolo quadrato di cartone nero che si muove su uno sfondo bianco a una velocità angolare costante, descrivendo una circonferenza attorno all'animale, a una distanza di sette centimetri. Il rospo interpreta uno stimolo di questo tipo come « preda » e cerca, per mezzo di successivi movimenti di rotazione, di mantenere l'oggetto fisso al centro del suo campo visivo. La misura dell'attività di orientamento è data dal numero delle risposte di rotazione al minuto.

L'ampiezza angolare dello stimolo – cioè l'angolo che esso sottende – influenza l'attività di orientamento (*si veda la figura a pagina 26*). Tra vari oggetti quadrati, i rospi preferiscono quelli il cui lato sottende un angolo di 4-8 gradi. (La grandezza assoluta di questi stimoli varia tra i 5 e i 10 millimetri. Esperimenti in cui viene variata la distanza tra lo stimolo e l'animale hanno dimostrato che ciò che conta è la grandezza assoluta e non quella angolare; nel comportamento di cattura i rospi mostrano una « costanza di dimensione ».) I rospi si allontanano da oggetti che sottendono angoli maggiori di 30 gradi mostrando una risposta di fuga. Dati più precisi si ottengono sostituendo agli stimoli quadrati sbarre di varia lunghezza. Quando uno stimolo di due gradi per due viene prolungato lungo l'asse orizzontale, l'attività di orientamento aumenta finché non viene raggiunto un livello di saturazione; gli oggetti a forma di verme ri-

Modelli di comportamento caratteristici del rospo *Bufo bufo*. L'attività viene di norma indotta nell'animale dalla percezione di stimoli ottici. Questi disegni, tuttavia, sono ricavati da fotografie di animali il cui cervello era stimolato elettricamente nel quadro delle ricerche dell'autore sulle basi nervose del comportamento guidato dalla vista. L'elettrodo che si vede sulla testa del rospo penetra nel cervello. Una corrente elettrica applicata al tetto ottico, un centro della visione nel cervello, induce una risposta di cattura della preda: orientamento o rotazione del corpo (*a*), schiocchi con la lingua (*b*) e pulizia della bocca (*c*). La stimolazione elettrica, invece, di un punto nel talamo sinistro o destro induce l'animale a una posizione difensiva (*d*, *e*), e la stimolazione di un'altra parte del talamo lo spinge ad accovacciarsi in una risposta di evitamento (*f*).

Analisi quantitativa delle risposte comportamentali del rospo a oggetti di varia forma e misura. Piccoli oggetti neri erano fatti muovere attraverso il campo visivo a una distanza di sette centimetri, e la risposta di orientamento era misurata nei rospi normali (*a sinistra*) e in altri cui era stato asportato il talamo (*a destra*). Nei rospi normali le risposte di cattura della preda (rotazione in direzione dell'oggetto) venivano indotte con maggiore efficacia da quadrati i cui lati sottendevano angoli di 4-8 gradi; quando i quadrati erano piú grandi, i rospi se ne allontanavano. Le sbarre verticali non erano efficaci come oggetto preda, e la loro efficacia era inversamente proporzionale all'altezza. Le sbarre orizzontali, a forma di verme, erano tanto piú efficaci come oggetto preda quanto piú erano lunghe, fino a un dato limite. Le sbarre doppie erano meno attraenti, e l'effetto variava con la distanza fra le due sbarre. Viene illustrato (*in basso*) il rapporto esistente fra l'effetto delle sbarre doppie e quello della sbarra singola. Nei rospi privati del talamo la risposta di orientamento è « disinibita ».

sultano essere particolarmente attraenti per i rospi. La risposta invece diminuisce quando un piccolo stimolo viene prolungato in senso verticale o perpendicolarmente rispetto alla direzione del movimento.

Altri esperimenti indicano che i rospi effettuano la discriminazione fra oggetto « preda » e oggetto « nemico » in base a configurazioni di punti o linee, considerando anche la direzione del movimento. Una catena orizzontale, costituita da diversi elementi di due gradi per due che si muovono nello stesso senso, ha il significato di preda. Uno di questi elementi che si muove da solo costituisce uno stimolo preda di grandezza immediatamente superiore alla soglia di risposta. Quando alla catena orizzontale viene associata una formazione verticale separata (per cui di fatto si realizza una struttura a forma di « L » che si muove lungo il suo lato piú lungo), essa perde efficacia in quanto stimolo al comportamento di cattura. L'effetto inibitore della formazione verticale dipende dalla sua distanza dall'elemento orizzontale. Se viene introdotta una seconda formazione verticale che dà luogo a una struttura a forma, piú o meno, di « U », la figura nel suo insieme assume il significato di nemico. L'interpretazione etologica è che questa forma rappresenta una « moltitudine », e nel cervello del rospo le interazioni inibitorie prima frenano il comportamento di cattura, e poi inducono il comportamento di fuga.

A forma e velocità angolare costanti, l'attività comportamentale in genere aumenta in proporzione diretta al grado di contrasto fra lo stimolo e lo sfondo. Gli oggetti bianchi che si muovono su sfondo nero sono di norma piú attraenti, come preda, che non oggetti neri su sfondo bianco; questi ultimi, d'altra parte, sono piú efficaci nell'indurre il comportamento di fuga. A grandezza e contrasto costanti, l'attività comportamentale aumenta con lo aumentare della velocità angolare, sino a un massimo tra 20 e 30 gradi al secondo. Gli oggetti immobili in genere non suscitano risposta né di cattura né di fuga. Il carattere comune avente valore determinante affinché uno stimolo assuma il significato di preda o di nemico è il movimento e la discriminazione fra i due tipi di stimolo viene effettuata soprattutto sulla base della loro forma: un oggetto esteso in senso orizzontale e parallelo alla direzione del movimento generalmente significa « preda », mentre un oggetto esteso perpendicolarmente alla direzione del movimento significa « non-preda » o « nemico ».

Registrazione effettuata per mezzo di elettrodi delle risposte fornite dai neuroni del rospo agli stessi oggetti della figura precedente. Gli elettrodi registravano gli impulsi nelle terminazioni situate nel tetto ottico di fibre provenienti da singole cellule gangliari, cellule retiniche su cui convergono, attraverso cellule intermedie, i segnali provenienti dai recettori (*a*). Ogni cellula gangliare ha un campo recettivo eccitatorio circondato da un campo recettivo inibitorio. Il diametro dei campi eccitatori e la forza della zona inibitoria circostante sono diversi per ciascuna delle tre classi di cellule gangliari (*b*). Per l'oggetto quadrato e la sbarra verticale (stimoli che vengono « confusi » dalla cellula gangliare) il massimo dell'attività viene indotto quando la grandezza dell'oggetto corrisponde alla grandezza del campo eccitatorio del rispettivo tipo di cellula gangliare (*c*). La lunghezza in senso orizzontale non influisce molto su questo tipo di risposta cellulare. L'aggiunta di una formazione verticale a una sbarra orizzontale ha, su queste cellule, un effetto minore che non sul comportamento (*si veda la figura nella pagina a fronte*).

Che tipo di messaggio viene trasmesso dall'occhio al cervello del rospo? Questo quesito venne formulato dapprima per la rana; esso costituí l'argomento di un'affascinante ricerca di Jerome Y. Lettvin e dei suoi colleghi al Massachusetts Institute of Technology, e in seguito venne studiato dal punto di vista quantitativo da O.J. Grüsser e dai suoi collaboratori alla Libera Università di Berlino. Porsi questo quesito significa aprire la « scatola nera » del cervello del rospo, o per lo meno analizzare le funzioni cerebrali che partecipano alla trasformazione delle afferenze prodotte dagli stimoli ottici nei relativi modelli comportamentali. Adesso descriverò i reperti neurofisiologici che ci permettono di stabilire se la codificazione degli stimoli-chiave « preda » e « nemico » ha luogo a livello della retina.

Nella retina del rospo vi sono tre tipi di cellule gangliari che inviano le loro fibre, per mezzo del nervo ottico, alla struttura mesencefalica chiamata tetto ottico. Introducendo un microelettrodo nel tetto è possibile registrare i potenziali d'azione, o segnali nervosi, dalle terminazioni di queste fibre. John E. Dowling, allora alla facoltà di medicina dell'Università Johns Hopkins, dimostrò, per mezzo della micrografia elettronica, che nella retina della rana, o del rospo, ogni cellula gangliare è connessa a un certo numero di recettori per mezzo di cellule bipolari e di cellule amacrine. Ogni cellula gangliare riceve cosí informazioni da una determinata regione del campo visivo dell'animale. Le connessioni laterali, formate dalle cellule orizzontali e amacrine, svolgono una funzione nel determinare le proprietà di questo campo recettivo. Nei rospi, come nelle rane, il campo è costituito da un settore circolare centrale eccitatorio immediatamente circondato da un campo recettivo inibitorio. Il movimento di un oggetto attraverso il campo eccitatorio induce una scarica nella cellula gangliare che viene inibita se, contemporaneamente, un oggetto si muove nel campo inibitorio. I tre tipi di cellule gangliari nel rospo (come nella rana) differiscono per molte caratteristiche, compreso, in particolare, il diametro del loro campo recettivo eccitatorio: questo è di circa 4 gradi per le cosiddette cellule gangliari della classe II, di circa 8 gradi

per quelle della classe III, di 12-15 gradi per quelle della classe IV. (La prima classe è stata individuata nelle rane ma non nei rospi.)

Con i microelettrodi noi misuriamo la frequenza di scarica della cellula gangliare per vedere in quale misura essa viene modificata quando gli oggetti (corrispondenti a quelli sopra descritti per gli esperimenti sul comportamento) vengono fatti muovere nel campo recettivo della cellula. La frequenza di scarica aumenta con l'aumentare della lunghezza del lato di un oggetto quadrato sino a che tale lunghezza non raggiunge all'incirca quella del diametro del campo eccitatorio; a questo punto, non appena l'oggetto diviene abbastanza grande da influenzare parte del circostante campo inibitorio, la frequenza decresce. Per ogni tipo di cellula, dunque, il massimo dell'attività è provocato da oggetti di diversa grandezza, in rapporto alla diversa grandezza dei campi eccitatori (*si veda l'illustrazione a pagina 27*). L'estendere un piccolo quadrato in senso orizzontale (facendolo diventare un « verme ») non modifica in alcun modo l'attività della cellula nervosa; questo è in netto contrasto con l'effetto prodotto sulla risposta comportamentale che abbiamo prima osservato. La dipendenza dell'attivazione nervosa dalla grandezza di uno stimolo è invece soprattutto funzione dell'estensione in senso perpendicolare alla direzione del movimento. Pertanto la frequenza di scarica prodotta da una stretta sbarra è quasi esattamente la stessa di quella prodotta da un quadrato che abbia la stessa altezza della sbarra. Una cellula gangliare della retina « confonde » i due stimoli, ma il rospo no; il quadrato eccita l'attività comportamentale mentre la sbarra la inibisce. Quando la grandezza dell'oggetto viene mantenuta costante, tuttavia, la dipendenza della frequenza di scarica dal grado di contrasto tra stimolo e sfondo e dalla velocità angolare è la stessa che si riscontra negli esperimenti sul comportamento.

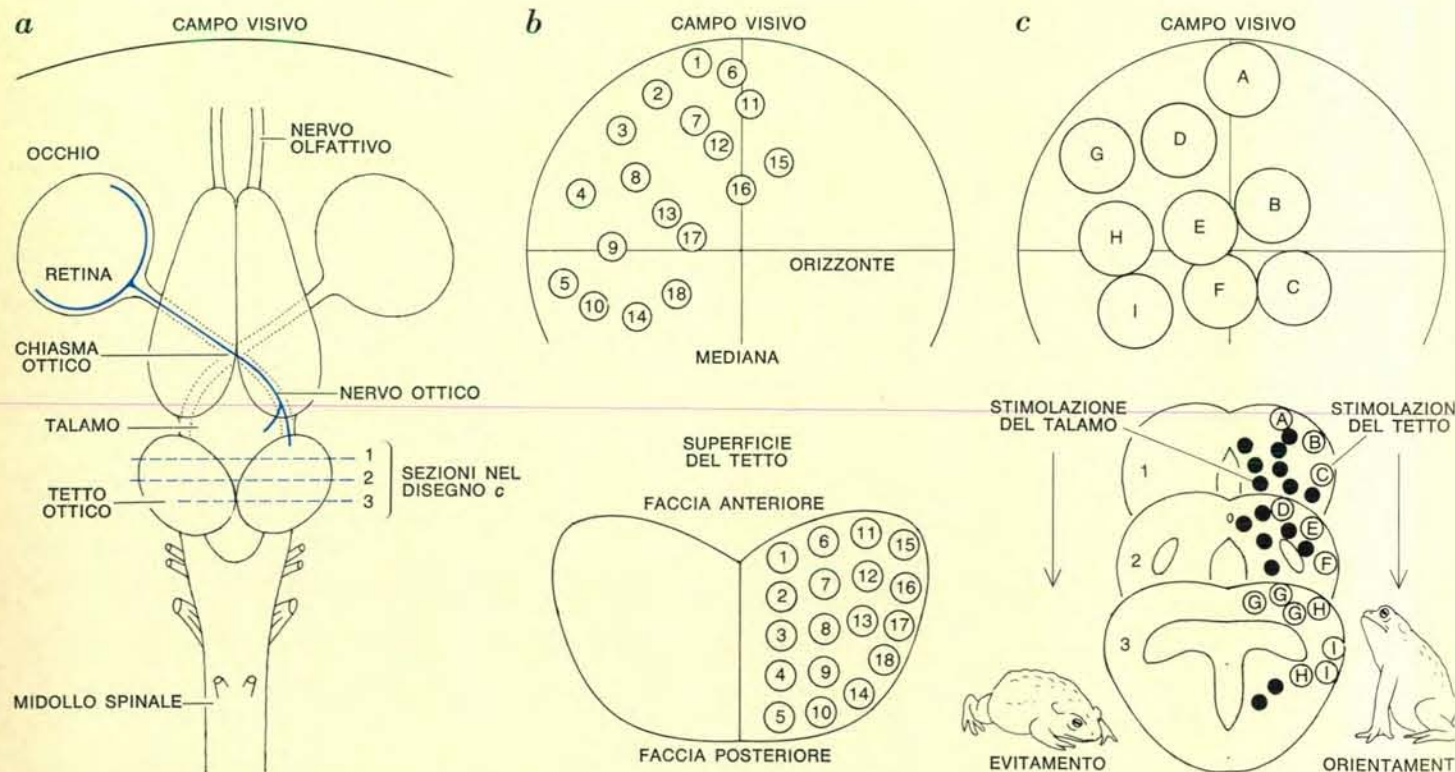
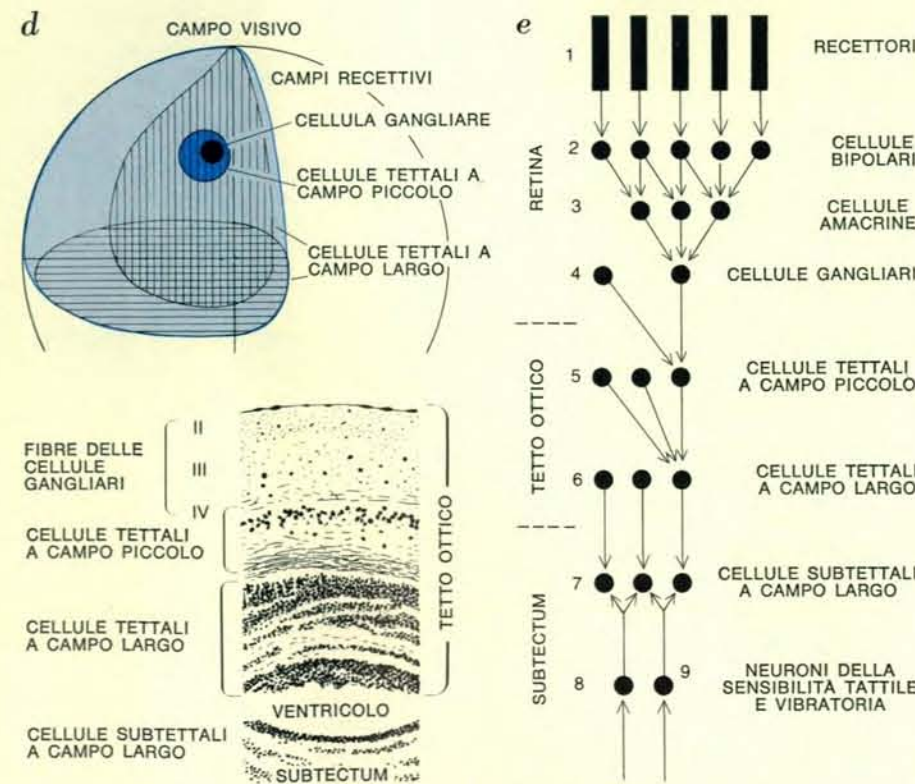
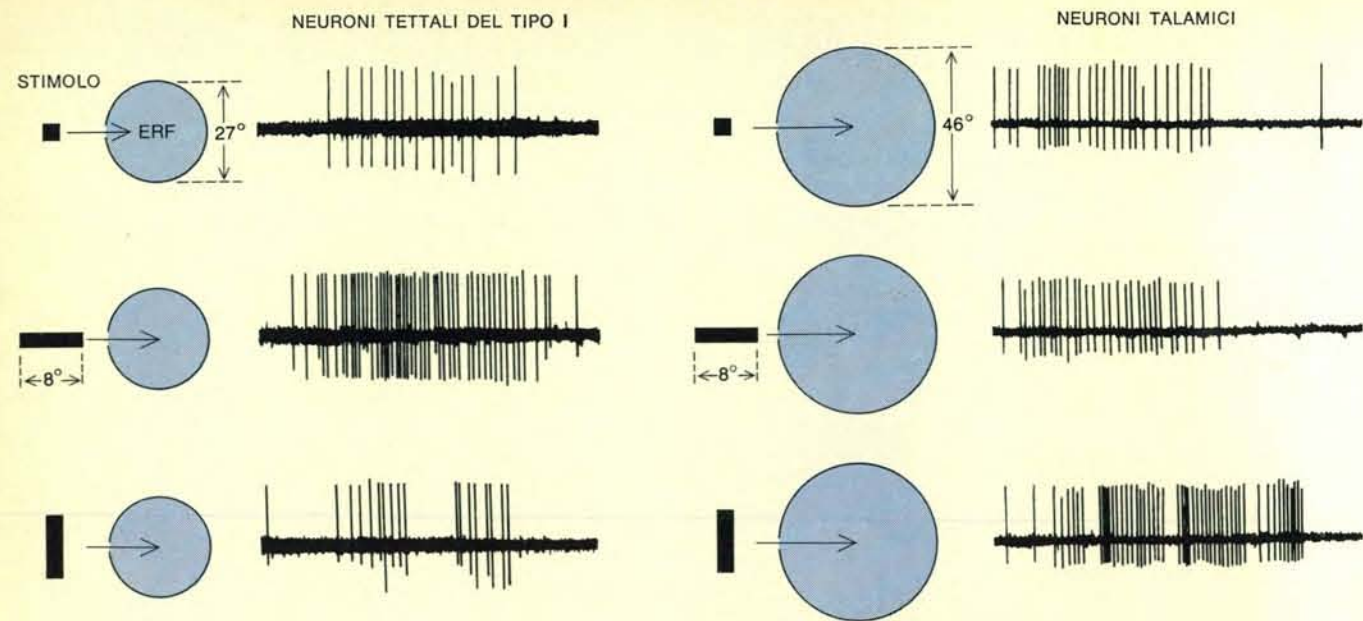
In sintesi, è chiaro che le prime importanti operazioni sulla informazione visiva trasmessa da uno stimolo preda o da uno stimolo fonte di paura sono effettuate dalla retina del rospo. Per ogni particolare stimolo preda o nemico, la risposta comportamentale alla velocità e al contrasto con lo sfondo sembra dipendere dalla elaborazione dell'informazione a livello retinico. I processi di inibizione e di eccitazione in rapporto alla grandezza, che sono stati osservati negli esperimenti sul comportamento e che svolgono una funzione essenziale nella discriminazione delle forme, non possono invece essere attribuiti all'influenza dei campi eccitatorio e inibitorio delle cellule gangliari retiniche. Nella retina non esistono « rivelatori di vermi » contrapposti a « rivelatori di nemici ». L'analisi differenziale, e dunque l'interpretazione dello stimolo che è rilevante ai fini del comportamento, deve essere effettuata dalle popolazioni di cellule nervose al di là del livello retinico.

Poiché caratteristiche differenti sono codificate da tutti i tipi di cellula gangliare, il quesito va posto nei seguenti termini: dove viene interpretata questa codificazione? Chi comunica al sistema nervoso centrale se un aumento della frequenza di scarica delle cellule gangliari deriva, per esempio, dall'aumento del contrasto fra stimolo e sfondo o da una maggiore grandezza dello stimolo? La differenziazione può essere fatta solo se gruppi separati di cellule ricevono afferenze diverse per mezzo di fibre differenti del nervo ottico. In realtà è proprio così. Le fibre del nervo ottico provenienti da ciascun occhio passano, attraversando il chiasma ottico, alla metà controlaterale del cervello e terminano in varie parti del telencefalo e del mesencefalo. Due di queste stazioni terminali sono particolarmente interessanti in questo contesto. Una, alla quale si proietta la maggior parte delle fibre del nervo ottico, si trova negli strati superficiali del tetto ottico nel mesencefalo. L'altra si trova nel talamo e nella regione pretettale del diencefalo.

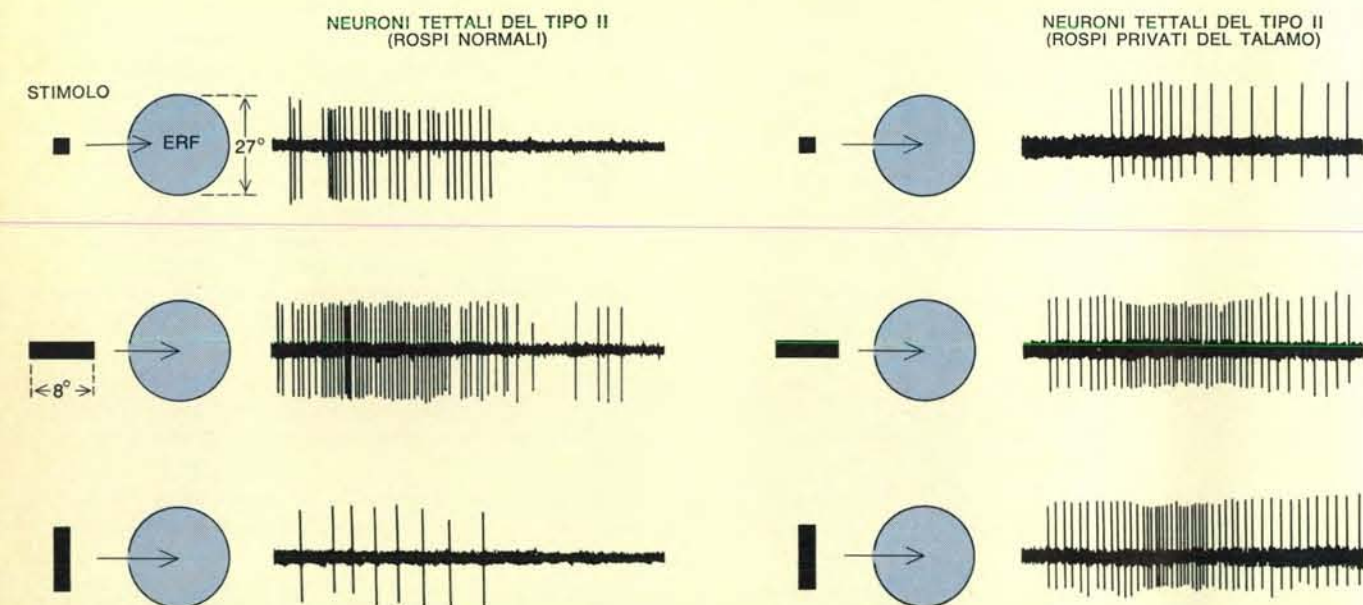
Il tetto ottico costituisce un sistema di localizzazione. Nel tetto esiste una mappa topografica precisa della retina e pertanto dell'intero campo visivo. Il movimento di un oggetto in una data parte del campo visivo eccita una corrispondente regione del tetto dove terminano le relative fibre del nervo ottico (*si veda l'illustrazione in queste due pagine*). La registrazione dell'attività elettrica di singoli neuroni, o cellule nervose, del tetto, dice in qual misura reagiscono le singole cellule gangliari della retina che li eccitano. In alcuni strati, per esempio, vi sono neuroni del tetto con un campo recettivo eccitatorio di circa 10-27 gradi che sono attivati esclusivamente da oggetti in movimento. Questi neuroni, probabilmente, rappresentano un sistema di localizzazione. Questa ipotesi è avvalorata da esperimenti nei quali noi stimoliamo la regione del tetto di rospi, che si muovono liberamente, con una serie di impulsi somministrati per mezzo di un elettrodo impiantato nel tetto. La stimolazione di una data parte del tetto fa sempre girare i rospi in direzione di una data parte del campo visivo. Presumibilmente i neuroni, che in tal modo noi attiviamo, hanno una connessione diretta con il sistema motorio dell'animale poiché (diversamente dai movimenti naturali di orientamento effettuati in risposta a un oggetto preda) la reazione di orientamento indotta elettricamente non è modificata dalla presentazione simultanea di un oggetto nemico.

Se l'elettrodo registratore viene introdotto piú in profondità nel tetto, esso incontra neuroni che hanno campi recettivi piú ampi. Alcuni di questi coprono l'intero campo visivo del lato opposto, altri tutta la parte bassa del campo, altri tutto il campo direttamente di fronte all'animale. Cosa abbastanza interessante, tutti e tre i tipi di neuroni includono il punto di fissazione, cioè il punto di massima acuità visiva vicino al centro del campo. Il grado di attivazione di questi tre tipi di neuroni potrebbe fornire al rospo delle informazioni circa la localizzazione di un oggetto voluminoso, poiché ogni volta che tutti e tre sono eccitati simultaneamente, l'oggetto deve trovarsi nel punto di fissazione.

Nelle situazioni naturali il comportamento dei rospi può essere influenzato da altre modalità sensoriali oltre la visione. Se, per esempio, uno scarafaggio attraversa il campo visivo, la reazione di orientamento del rospo può essere accelerata o ritardata da simultanei stimoli vibratori e tattili. Risultati di questo genere si possono ottenere sperimentalmente presentando agli animali dei modelli preda insieme a stimoli acustici o tattili. L'area in cui si producono tali modificazioni dell'attività comportamentale sembra essere localizzata nella regione subtettale, dove si realizza l'integrazione multisensoriale. Nell'area sottostante al terzo ventricolo del mesencefalo vi sono neuroni aventi un ampio campo recettivo simile a quello delle cellule degli strati profondi del tetto. Questi neuroni subtettali ricevono afferenze addizionali da neuroni eccitati da stimoli tattili e vibratori. Il campo di recezione « meccanica » di uno di questi neuroni bimodali è sempre localizzato dalla stessa parte del campo recettivo visivo. Le afferenze addizionali di questi neuroni possono servire ad abbassare la soglia di una parte del campo visivo in cui si prevede la presentazione di uno stimolo ottico, realizzando in questo modo un aumento del livello di vigilanza visuale.

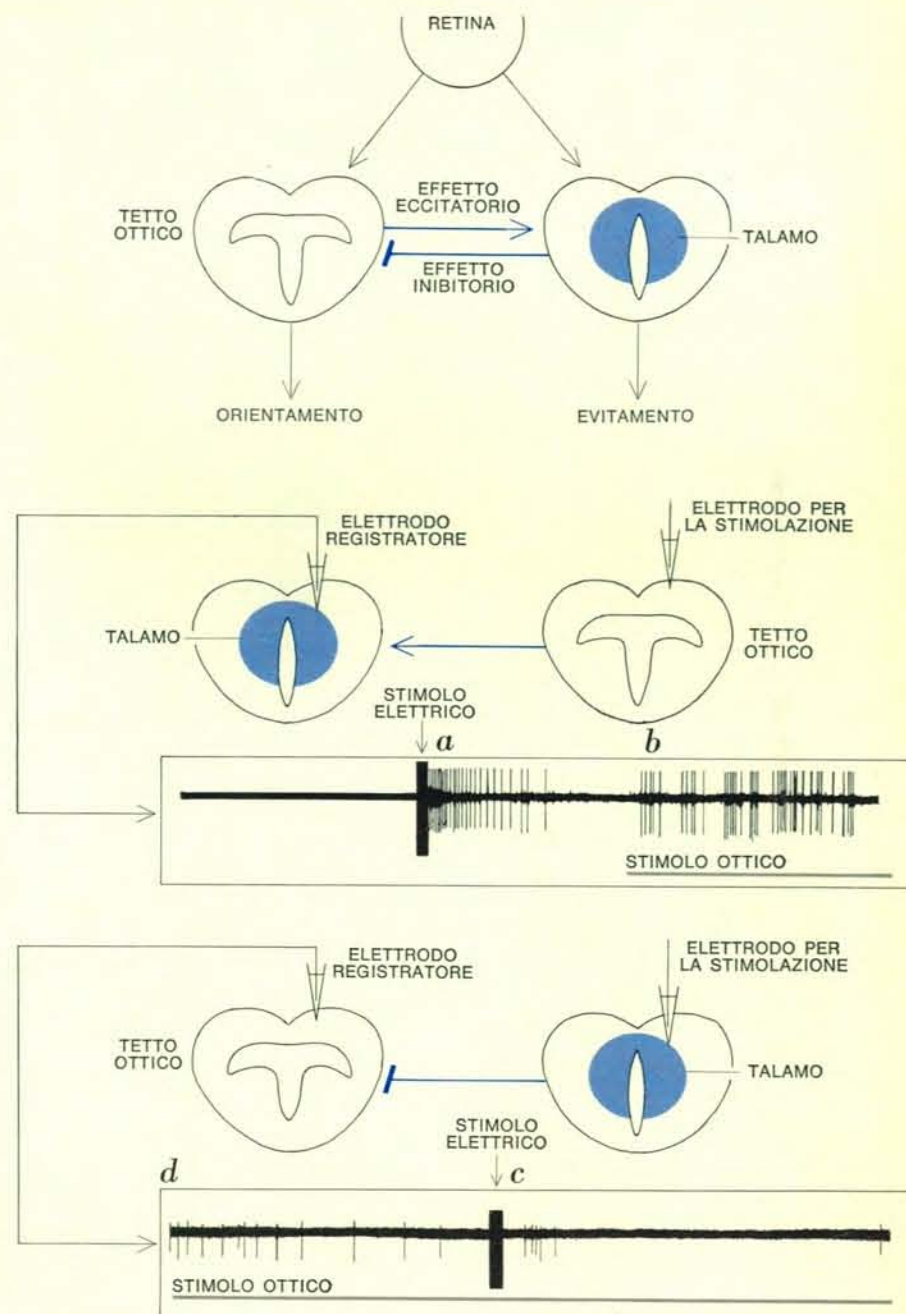
Il tetto ottico comprende anche un sistema di neuroni che elabora alcuni aspetti degli stimoli in movimento di importanza rilevante per il comportamento (*si vedano le illustrazioni nella pagina seguente*). Le cellule hanno campi recettivi eccitatori del diametro di circa 27 gradi. Quelle designate come neuroni tettali del tipo I vengono attivate soprattutto se la superficie dello stimolo, costituito da un oggetto che si muove nel campo recettivo, si estende in senso parallelo alla direzione del movimento; l'estensione in senso perpendicolare alla direzione del movimento non produce lo stesso effetto. Altre cellule, i neuroni tettali del tipo II, differiscono dai precedenti perché la loro frequenza di scarica, invece, diminuisce in proporzione diretta all'estensione della superficie dello stimolo in senso perpendicolare alla direzione del movimento. La risposta di questi neu-

Gli esperimenti di neurofisiologia forniscono dei dati sulle funzioni di diverse parti del sistema ottico del rospo. Le fibre delle cellule gangliari della retina si proiettano prevalentemente al tetto ottico e al talamo (*a*). Il campo visivo di ciascun occhio è mappato (*b*) sulla superficie dorsale della metà controlaterale del tetto sulla base di un rapporto 1 a 1 (*numeri*). Analogamente la stimolazione elettrica sperimentale (*c*) di varie parti del tetto (*lettere*) induce il rospo a girarsi in direzione delle parti corrispondenti del campo visivo. D'altra parte la stimolazione (*cerchietti neri*) del talamo, che è situato in parte sotto il tetto, induce l'azione opposta, e cioè una reazione di avversione o allontanamento. Quando l'elettrodo registratore penetra in profondità nel tetto, esso incontra successive popolazioni di cellule con differenti campi recettivi (*d*). Vi sono cellule a campo piccolo con campi appena piú grandi di quello delle cellule gangliari e, piú in profondità, tre tipi di cellule a campo largo, ciascuna avente un raggio diverso (*in colore, tratteggio orizzontale e tratteggio verticale*). Un disegno basato su una sezione colorata di cervello indica gli strati in cui queste cellule rispettivamente si trovano. L'ultimo disegno (*e*) pone in rapporto le varie popolazioni di cellule fra loro e mostra un altro strato di cellule a campo largo che ricevono afferenze da cellule visive sovrastanti e anche da cellule che rispondono a stimoli tattili o a stimoli vibratori.

L'identificazione dei caratteri dello stimolo oltre il livello retinico viene realizzata da cellule del tetto e del talamo. La registrazione dell'attività nervosa delle singole cellule mostra che i neuroni tettali del tipo I (*a sinistra*) sono attivati prevalentemente da oggetti che si muovono nel campo visivo e il cui asse più lungo è parallelo alla direzione del movimento. Le cellule dell'area talamica (*a destra*) rispondono prevalentemente a oggetti estesi perpendicolarmente alla direzione del movimento.

roni costituisce lo stimolo-chiave perché un oggetto venga identificato come preda. Essi possono, cioè, essere considerati come il sistema di innesco della risposta di cattura.

La regione talamico-pretettale, la seconda stazione terminale delle fibre provenienti dalla retina, sembra fornire quello che può esser chiamato un sistema di « guardia ». Recentemente, per mezzo di registrazioni a livello di singole cellule, ho identificato quattro tipi fondamentali di neuroni sensoriali visivi nel talamo del rospo. Essi sono attivati rispettivamente da quattro distinte situazioni-stimolo: 1) movimento di oggetti fonte di paura estesi perpendicolarmente alla direzione del movimento; campo recettivo eccitatorio di circa 46 gradi; 2) movimento di un oggetto in direzione del rospo; campo di circa 90 gradi; 3) grandi oggetti immobili; campo di circa 45 gradi; 4) stimolazione dei recettori dell'equilibrio nell'orecchio del rospo in seguito a variazioni di posizione. In generale

Le unità che fanno scattare l'intera risposta di cattura della preda sono probabilmente i neuroni tettali del tipo II. Nel rospo normale (*a sinistra*) le cellule sono attivate soprattutto da oggetti a forma di verme (*sbarra orizzontale*). Esse sono attivate in misura minore (e la diminuzione è più marcata che nel caso dei neuroni tettali del tipo I) da stimoli che negli esperimenti sul comportamento sono irrilevanti alla reazione di cattura (*sbarra verticale*). Dopo asportazione del talamo, tuttavia (*a destra*), la loro risposta a questi stimoli irrilevanti è molto aumentata; sembra quindi che il segnale talamico sia inibitorio.

questi neuroni talamici vengono attivati principalmente in situazioni che tendono a provocare movimenti di difesa-allontanamento da un nemico, spostamenti in senso laterale o compenso a posizioni di perdita dell'equilibrio corporeo. Esperimenti di stimolazione cerebrale convalidano la nostra impressione che nella regione talamico-pretettale possano confluire le reazioni che portano a movimenti di tipo autoprotettivo. La stimolazione elettrica di varie parti della regione provoca le seguenti reazioni: gli animali chiudono la rima palpebrale o fanno il gesto di tuffarsi, si rivoltano dall'altra parte, schizzano via con manifestazioni di panico o inclinano il corpo.

La nostra ipotesi di lavoro prevede l'esistenza di connessioni tra il tetto ottico e la regione talamico-pretettale; le stimolazioni elettriche nel tetto provocano soprattutto reazioni di orientamento mentre le stimolazioni nella regione talamico-pretettale inducono il comportamento di evitamento. In situazione naturale, piccole prede a forma di verme fanno partire impulsi che vanno a particolari strati del tetto. Oggetti voluminosi estesi perpendicolarmente alla direzione del movimento stimolano particolari neuroni nella regione talamico-pretettale, sia direttamente attraverso afferenze retiniche sia indirettamente attraverso il tetto ottico. Questi neuroni talamico-pretettali inibiscono, a loro volta, il tetto e possono anche attivare il comportamento di evitamento (*si veda l'illustrazione in questa pagina*).

L'esistenza delle connessioni postulate tra le strutture del mesencefalo e del diencefalo è stata dimostrata fisiologicamente in due modi. Un modo è rappresentato dalla stimolazione elettrica diretta. I neuroni talamici che sono sensibili al movimento possono essere attivati anche dalla stimolazione di determinati punti del tetto ottico. Quando si scambiano gli elettrodi stimolanti e quelli registratori, la risposta dei neuroni tettali del tipo II a oggetti in movimento può essere inibita dalla stimolazione di cellule nel talamo. Il secondo modo è rappresentato da un intervento chirurgico: se viene asportato il tetto ottico, i movimenti di orientamento vanno perduti, e così le reazioni di evitamento, segno chiaro dell'esistenza di vie nervose dal tetto al talamo. Se si asporta la regione talamico-pretettale senza danneggiare il tetto, va perduto il comportamento di evitamento e la risposta di orientamento viene disinibita in modo spettacolare, anche in presenza di oggetti nemici; questa può essere considerata la prova dell'esistenza di vie inibitorie dal talamo al tetto ottico. Nei rospi privati della regione talamico-pretettale, la disinibizione si estende a tutto il campo visivo della metà controlaterale; piccole lesioni nella regione talamico-pretettale colpiscono solo piccoli settori localizzati del campo visivo. Esperimenti quantitativi effettuati su rospi mancanti della regione talamico-pretettale dimostrano chiaramente che questi animali non sanno discriminare fra stimoli che sono rilevanti per il comportamento e stimoli che non lo sono. La risposta delle cellule tettali del tipo II a stimoli in movimento mostra un analogo effetto di « disinibizione » dopo l'asportazione della regione talamico-pretettale (*si vedano le illustrazioni a pagina 26 e a pagina 30 in basso*).

Gli esperimenti precedenti hanno messo in evidenza la presenza di connessioni fra il tetto ottico e il talamo; i segnali provenienti dalla retina eccitano sia il tetto sia il talamo; successivamente, impulsi provenienti dai neuroni tettali del tipo I eccitano ulteriormente le cellule del talamo, mentre i segnali provenienti dal talamo inibiscono l'attività dei neuroni tettali del tipo II. Le due strutture danno quindi il via a due diversi tipi di attività motoria (*in alto*). Una ulteriore conferma della presenza di queste connessioni è stata fornita da esperimenti di stimolazione elettrica. La stimolazione del tetto (*al centro*) fa partire impulsi (*a*) da cellule del talamo che di norma rispondono a stimoli ottici (*b*). La stimolazione del talamo (*in basso*) inibisce (*c*) gli impulsi che di solito sono indotti (*d*) dagli oggetti in movimento nei neuroni del tipo II.

I risultati sperimentali sinora descritti fanno ritenere che la sequenza degli eventi sia la seguente: sulla base delle informazioni trasmesse dalle cellule gangliari della retina, il tetto ottico comunica al rospo in quale punto del campo visivo è situato uno stimolo, quanto è grande, quanto intenso è il suo contrasto con lo sfondo e quanto rapidamente si muove. Le connessioni esistenti fra il tetto ottico e le strutture della regione talamico-pretettale permettono al rospo di discernere il significato che tali segnali visivi hanno in relazione al suo comportamento. Il fondamentale processo di filtraggio per la differenziazione fra preda e nemico può essere concepito sotto forma di passaggio attraverso una serie di « finestre di discriminazione » (*si veda l'illustrazione in questa pagina*), in cui ciascuno stadio analizza un particolare aspetto dell'oggetto in questione.

Ogni cellula gangliare della retina funziona come una finestra verticale che codifica l'estensione in senso perpendicolare alla direzione del movimento.

L'analisi retinica è ripetuta e amplificata nella regione talamico-pretettale, dove un gruppo di neuroni svolge anch'esso la funzione di finestra verticale, questa volta con una certa soglia di risposta minima. L'estensione in senso orizzontale viene codificata innanzi tutto dalle cellule tettali del tipo I che formano una finestra orizzontale. Le cellule tettali del tipo II eseguono una somma fra i segnali aventi effetto inibitorio che arrivano dalla regione talamico-pretettale e quelli aventi effetto eccitatorio provenienti dalle cellule del tipo I. Il segnale risultante agisce come stimolo di innesco della reazione di orientamento. Il comportamento di evitamento viene probabilmente realizzato attraverso l'attivazione di un altro gruppo di neuroni talamico-pretettali, e l'attivazione è proporzionale all'apporto di afferenze supplementari provenienti da due dei gruppi di cellule con funzione di « finestra di discriminazione ».

Uno degli aspetti notevoli di questo sistema è un certo grado di plasticità o modificabilità. Durante i mesi estivi gli oggetti preda bianchi fatti muovere su uno sfondo nero suscitano il comportamento di orientamento in modo molto piú efficiente che non oggetti neri su uno sfondo bianco. In autunno e in inverno la situazione è inversa, e allo stesso tempo l'attività di cattura dei rospi va incontro a una diminuzione globale. Recentemente i nostri elettrodi registratori hanno rilevato che l'attività di singole cellule gangliari della classe II della retina mostra la medesima variazione stagionale nella preferenza bianco-nero. D'inverno i neuroni con campo recettivo nella parte bassa del campo visivo sono attivati di piú da oggetti neri che da oggetti bianchi; nel campo superiore la situazione è inversa. D'estate invece, i neuroni, il cui campo recettivo si trova nella metà superiore del campo visivo, sono attivati soprattutto dagli stimoli neri, mentre i neuroni recettivi alla metà inferiore del campo visivo divengono piú sensibili agli stimoli bianchi, e cosí rimangono finché in autunno non divengono nuovamente dominanti gli stimoli neri.

Qual è il significato biologico di queste osservazioni? Si può osservare che per i rospi, la cui attività si svolge prevalentemente al crepuscolo, gli stimoli preda biologicamente importanti che compaiono nella metà inferiore del campo visivo sono piú chiari dello sfondo su cui si muovono; quelli che compaiono nella metà superiore del campo, invece, sono per la maggior parte relativamente scuri, o per lo meno altrettanto scuri che chiari. Queste relazioni di contrasto potrebbero riflettersi rispettivamente nelle caratteristiche di sensibilità delle cellule gangliari retiniche della classe II. Con l'approssimarsi dell'inverno e del periodo di ibernazione i rospi interrompono l'attività di cattura. Che cosa provoca questa interruzione? Un meccanismo potrebbe essere rappresentato dall'inversione delle caratteristiche di risposta delle cellule gangliari a opera di segnali trasmessi dal cervello alla retina, con la conseguenza che i rapporti di contrasto fra stimolo e sfondo risultano sfasati rispetto al mondo reale e gli oggetti preda diventano meno visibili. Per il rospo, in altre parole, oggetti identici appaiono diversi se visti d'inverno anziché d'estate, il che sta a ricordarci come l'immagine che un organismo si fa dell'ambiente sia un prodotto della sua attività cerebrale.

L'identificazione di un oggetto come « preda » o « nemico » viene raffigurata simbolicamente sotto forma di una serie di operazioni che hanno luogo a livello di « finestre di discriminazione » (*a*). Una cellula gangliare della retina codifica l'estensione in senso verticale (perpendicolare alla direzione del movimento), e in pratica risponde allo stimolo ottico nella misura in cui esso appare nella finestra verticale; la superficie che si estende oltre la finestra ha effetto inibitorio. Le cellule talamiche si comportano allo stesso modo. Nel tetto le cellule del tipo I codificano l'estensione orizzontale (in direzione del movimento). Le cellule tettali del tipo II sommano il segnale eccitatorio proveniente dalle cellule del tipo I con il segnale inibitorio proveniente dal talamo e il segnale risultante fa scattare la risposta di orientamento. A ogni stadio la frequenza di scarica della cellula dipende dal rapporto esistente fra l'oggetto e la finestra che percepisce la sua estensione in senso perpendicolare o in senso parallelo alla direzione del movimento (*b, c, d*). (La frequenza di scarica è qui rappresentata in forma schematica.)

In contrasto con la plasticità dei sistemi preposti al filtraggio e all'immagazzinamento delle informazioni e al riconoscimento delle forme, i meccanismi cerebrali implicati nei comportamenti istintivi sono molto rigidi e non si adattano facilmente a cambiamenti delle situazioni stimolo. Per ogni azione istintiva del repertorio comportamentale esiste un circuito nervoso programmato e « stampato » dalla cui coordinazione scaturisce l'appropriato atto motorio – anche se questo è divenuto inappropriato! Se un circuito del genere viene innescato (sia per via naturale sia per mezzo di una stimolazione elettrica) la reazione innata procede automaticamente. Il comportamento di cattura ne è un esempio dimostrativo. Sulla base di esperimenti di stimolazione cerebrale noi crediamo che la sequenza di eventi che controlla una naturale risposta di orientamento sia all'incirca la seguente: uno stimolo naturale determina la formazione di un'immagine su una porzione della retina al di fuori della regione di fissazione; il locus retinico ha un corrispondente locus di proiezione nel tetto ottico. Se il processo di filtraggio che abbiamo sopra descritto ha identificato l'oggetto come preda, viene attivato l'appropriato circuito nervoso. Il valore corrispondente alla distanza fra l'immagine della preda sulla retina e il punto di fissazione viene trasmesso al sistema motorio del rospo. Il risultato è la risposta di orientamento: un movimento di rotazione tale da portare la rappresentazione retinica della preda nel punto di fissazione. Questo invia un impulso corrispondente al locus del tetto ottico. Questo impulso innescante, non appena raggiunge il valore soglia, fa scattare la rimanente sequenza di cattura della preda, del tutto indipendentemente dal risultato o anche dal beneficio immediato derivante all'animale da questa attivazione. Per esempio, se un oggetto preda sperimentale viene rimosso nell'istante in cui esso è fissato dal rospo, la sequenza si verifica ugualmente per intero. Il rospo fa schiocchi con la lingua, deglutisce e si lecca la bocca nonostante il « vuoto situazionale ». La sequenza è simile, nella sua ineluttabilità, alla serie di eventi che si manifesta quando la regione scatenante del tetto viene stimolata per mezzo di un elettrodo.

Per quanto riguarda il comportamento di evitamento, i risultati della stimolazione talamica indicano che esso è controllato e diretto da un unico programma. La risposta del rospo consiste nel piantare fermamente a terra le zampe dello stesso lato del corpo e nel raccogliere quelle della parte opposta. Quando il rospo si trova immobile e in equilibrio in questa posizione, i moduli comportamentali supplementari per correggere l'inclinazione del corpo o effettuare i vari movimenti di difesa si susseguono automaticamente.

I risultati che ho esposto dimostrano che in un vertebrato inferiore i processi nervosi preposti alla localizzazione e all'identificazione di uno stimolo ottico e allo scatenamento delle risposte motorie istintive a esso associate sono topograficamente separati ma intimamente connessi fra loro. Nel corso dell'evoluzione i centri di due di questi processi, la localizzazione visiva e l'azione istintiva, sembrano essere rimasti approssimativamente nella loro posizione originaria. Essi occupano le stesse aree del cervello, il tetto e il talamo, nelle scimmie e nei gatti come nei rospi. L'organizzazione di queste parti del cervello, il cui studio è stato reso possibile dall'applicazione di metodi sia neurofisiologici sia etologici, mostra una notevole costanza in tutte le classi di vertebrati. Non cosí per l'identificazione dello stimolo. Nei rospi questo processo ha luogo fondamentalmente nella regione talamico-pretettale e anche nella retina e nel tetto. I mammiferi, invece, sono andati incontro a una ulteriore evoluzione in rapporto all'importanza che il riconoscimento della forma ha assunto nell'evoluzione del loro comportamento. Si è sviluppato un nuovo substrato per due processi associati, ma altamente specializzati, il filtraggio e l'immagazzinamento dell'informazione: la corteccia cerebrale. Dalle ricerche di Gerald E. Schneider compiute al Massachusetts Institute of Technology abbiamo appreso che, in questo caso, l'ontogenesi riflette la filogenesi. Nel criceto neonato la discriminazione della forma ha la sua base a livello delle vie sottocorticali fra il tetto e il talamo. Nell'animale adulto, invece, essa ha luogo nella corteccia.

FISICA

LE SCIENZE
edizione italiana di
SCIENTIFIC AMERICAN

ha finora pubblicato i seguenti articoli:

LA SUPERCONDUTTIVITÀ ALLE ALTE PRESSIONI
di N.B. Brandt e N.I. Ginzburg (n. 35)

PROGRESSI NEL RICONOSCIMENTO DELLE CONFIGURAZIONI
di R.G. Casey e G. Nagy (n. 35)

LA RIVELAZIONE DELLE ONDE GRAVITAZIONALI
di J. Weber (n. 36)

LA STRUTTURA DEL PROTONE E DEL NEUTRONE
di H.W. Kendall e W.K.H. Panofsky (n. 37)

LA FUSIONE CON IL LASER
di M.J. Lubin e A.P. Fraas (n. 37)

LE BOLLE MAGNETICHE
di A.H. Bobeck e H.E.D. Scovil (n. 37)

UNA NUOVA CLASSE DI LASER A DIODO
di M.B. Panish e I. Hayashi (n. 38)

FOTONI « PESANTI »
di F.V. Murphy e D.E. Yount (n. 38)

LA SPETTROSCOPIA MÖSSBAUER
di R.H. Herber (n. 41)

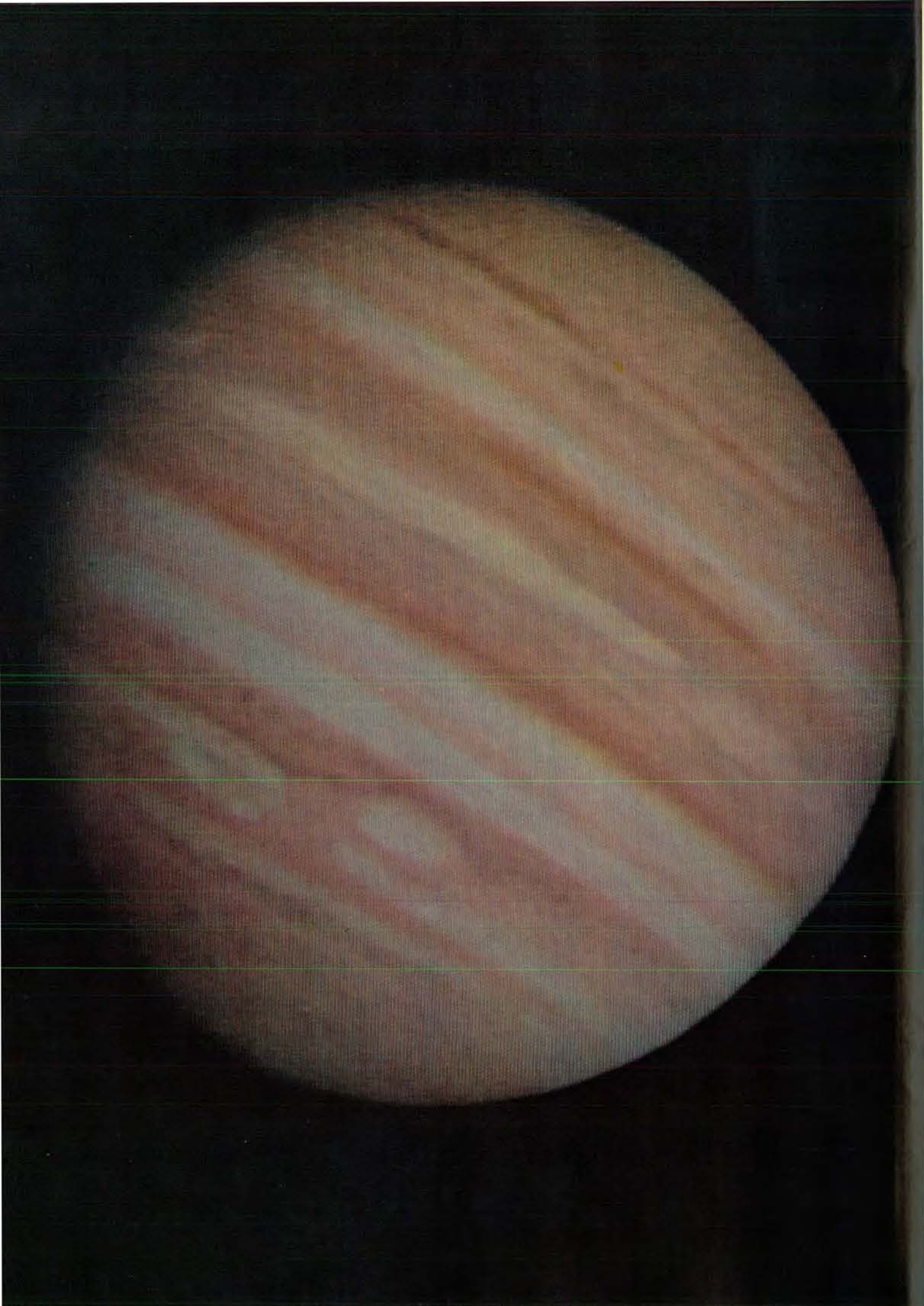
GLI ANELLI DI ACCUMULAZIONE
di B. Touschek (n. 42)

NUOVI SUPERCONDUTTORI
di T.H. Geballe (n. 42)

LA PRESSIONE DELLA LUCE LASER
di A. Ashkin (n. 45)

L'ELETTROSTATICA
di A.D. Moore (n. 46)

SUPERCONDUTTORI PER LA TRASMISSIONE DI ENERGIA ELETTRICA
di D.P. Snowden (n. 47)

IL TOKAMAK E LA FUSIONE TERMONUCLEARE CONTROLLATA
di B. Coppi e J. Rem (n. 50)

L'evoluzione chimica del sistema solare

I processi chimici e fisici che portarono alla formazione di tutti i corpi del sistema solare da una iniziale nube di polvere e gas

di John S. Lewis

L'era delle esplorazioni spaziali con l'atterraggio di astronavi sulla Luna e su Venere e le missioni interplanetarie verso Marte, Venere e Giove e ora anche verso Mercurio e Saturno ha allargato notevolmente la nostra conoscenza sulla composizione dei pianeti e dei satelliti e sull'evoluzione del sistema solare. La storia della moderna conoscenza della composizione del sistema solare cominciò negli anni trenta col lavoro di Rupert Wildt sulla fisica e la chimica dei sistemi gioviani: Giove, Saturno, Urano e Nettuno.

Intorno al 1950 vi fu un secondo periodo di estese ricerche sulla natura chimica del sistema solare. A quel tempo Harrison Brown propose che si potessero dividere gli oggetti che formano il sistema solare in tre classi secondo la loro densità e la loro composizione chimica. La prima classe era composta da oggetti rocciosi come i pianeti terrestri, i loro satelliti, i pianetini o asteroidi e i meteoroidi. Alla seconda classe appartenevano oggetti composti di materiali rocciosi e ghiaccio come i nuclei delle comete e i satelliti dei pianeti esterni. La terza classe era composta dal Sole e dai pianeti gioviani che sono costituiti principalmente di materia allo stato gassoso. Contemporaneamente al suggerimento di queste classi da parte di Brown, Gerard P. Kuiper e Harold C. Urey stavano lavorando su teorie complesse relative all'origine del sistema solare. Da allora lo studio particolareggiato delle rocce raccolte sulla Luna, gli spettri delle comete ottenuti di recente e le numerose fotografie della Luna e dei pianeti fatte dai veicoli spaziali, insieme a ciò che sappiamo sulla geologia della Terra, hanno contribuito globalmente al quadro attuale del sistema solare e della sua origine.

La nebulosa solare

Il sistema solare non è una miscela omogenea di elementi chimici. Allontanandosi dal Sole si riscontra una tendenza generale, per quanto concerne l'abbondanza degli elementi volatili cioè che evaporano facilmente, ad aumentare rispetto all'abbondanza degli elementi non volatili. Questa tendenza generale è chiaramente una prova a favore dell'ipotesi che i pianeti gioviani e gli altri corpi del sistema solare esterno si formarono a basse temperature e i pianeti terrestri si formarono a temperature alte. Secondo questa ipotesi il sistema solare si è formato dalla nebulosa solare: una vasta nube di polvere e di gas. Il centro della nebulosa era caldo (molte migliaia di gradi kelvin) mentre i bordi erano freddi (poche decine di gradi kelvin).

La differenza di temperatura significa che la composizione della polvere doveva variare radicalmente con la distanza dal centro. Vicino al protosole, al centro, tutti i materiali avrebbero dovuto essere completamente evaporati. A una distanza di forse trenta milioni di chilometri dal protosole, un quinto della distanza dell'orbita odierna della Terra, assai poca materia non volatile potrebbe essersi condensata in particelle solide. Da questo punto fino alla posizione attuale della fascia dei pianetini, la polvere dovrebbe essere stata composta soprattutto di grani di materia rocciosa con un contenuto solo limitato di sostanze volatili. Al di là della posizione della fascia degli asteroidi, la nebulosa sarebbe stata sufficientemente fredda da permettere alle sostanze volatili come acqua, ammoniaca e metano di congelare nei rispettivi ghiacci formando cosí particelle solide. Si ritiene che il Sole si sia formato quando il centro della nebulosa solare riscaldata divenne instabile e collassò per effetto della sua stessa gravità. I massicci pianeti gioviani potrebbero essersi formati mediante lo stesso processo. Oppure potrebbero aver avuto inizio con un nucleo roccioso e ghiacciato tanto grande che la sua attrazione gravitazionale avrebbe catturato grandi masse di materiale gassoso non differenziato. Su

Il pianeta Giove, visto in questa immagine dal *Pioneer 10*, è composto per lo piú dagli stessi elementi che compongono il Sole e nella stessa proporzione. È il prototipo di una delle tre classi di composizione di oggetti del sistema solare, quella comprendente il Sole e Saturno. Le altre due classi comprendono oggetti rocciosi come i pianeti terrestri e oggetti composti di materiali sia di roccia sia di ghiaccio, come i nuclei delle comete. In questa fotografia le fasce di nubi di Giove sono risolte con una ricchezza di particolari non rilevabili dalla Terra. Il pianeta appare gibboso a causa della posizione del veicolo spaziale rispetto al pianeta e al Sole, una veduta che non si può mai avere dalla Terra. Il terminatore, la linea tra l'emisfero illuminato e quello in ombra, è a sinistra sull'immagine. La fotografia è stata fatta a una distanza di 2 020 000 chilometri dal pianeta alle 20h 58min (tempo universale) il 2 dicembre 1973, circa 30 ore prima del punto di minima distanza, 130 000 chilometri, del *Pioneer* dal pianeta. È realizzata con l'esperimento del fotopolarimetro accoppiato a un intensificatore di immagini messo a punto dall'Università dell'Arizona. Il fotopolarimetro consiste di un telescopio con un'apertura di 25 cm accoppiato a un sistema ottico che divide la luce del pianeta in un canale blu e uno rosso. Il dispositivo spazzolava il pianeta mentre il *Pioneer* ruotava sul suo asse 4,8 volte al minuto, costruendo l'immagine in un reticolo in modo simile all'immagine che si forma su uno schermo televisivo. Il fotopolarimetro passava alternativamente dal canale blu a quello rosso ogni mezzo millesimo di secondo e ne telemetrava i dati grezzi a terra. All'Università dell'Arizona veniva rettificata l'immagine di ciascun canale per ridurre le distorsioni e aumentare la nitidezza, poi le due immagini venivano combinate in modo da ottenere la composizione a colori.

SOLE
MERCURIO
VENERE
TERRA
MARTE
GIOVE
SATURNO
URANO
NETTUNO
PLUTONE

OGGETTO	DISTANZA DAL SOLE (MILIONI DI CHILOMETRI)	DIAMETRO (MIGLIAIA DI CHILOMETRI)	MASSA (TERRA = 1)	NUMERO DI SATELLITI	DENSITÀ (ACQUA = 1)	PERIODO DI ROTAZIONE (GIORNI)	RIVOLUZIONE INTORNO AL SOLE (ANNI)	TEMPERATURA MEDIA (GRADI K)	GRAVITÀ (TERRA = 1)
SOLE	–	1395	343 000	–	1,4	27	–	5800	28
MERCURIO	58	4,8	0,1	0	5,4	55	0,24	~ 600	0,37
VENERE	108	12,2	0,8	0	5,1	–243	0,62	750	0,89
TERRA	150	12,7	1,0	1	5,5	1	1,00	180	1,00
MARTE	229	6,8	0,1	2	3,9	1	1,88	140	0,38
GIOVE	782	143	317,8	12	1,3	0,4	11,86	128	2,65
SATURNO	1435	122	95,2	10	0,7	0,4	29,48	105	1,14
URANO	2881	48	14,5	5	1,6	0,4	84,01	70	0,96
NETTUNO	4522	45	17,2	2	2,3	0,6	164,79	55	1,53
PLUTONE	6083	5,9	<0,2?	0	?	6,4	248,4	?	?

questa base si comprende facilmente l'analogia di composizione tra questi pianeti e il Sole.

La composizione dei pianeti terrestri e di oggetti analoghi varia con la distanza dal Sole. Esiste un modello concettualmente semplice, l'ipotesi della condensazione in equilibrio, che spiega queste differenze. Ammettiamo che, nella nebulosa solare, l'equilibrio chimico tra polvere e gas fosse responsabile della composizione del materiale solido che si è poi accumulato nei pianeti, nei loro satelliti, nei pianetini e nelle comete. Su questa base si può calcolare come la composizione chimica della materia condensata dipenda dalla sua temperatura al momento della formazione. A questo scopo è necessario un dato supplementare: le abbondanze relative dei vari elementi chimici di cui era composta la nebulosa primordiale.

Le nostre conoscenze sulle abbondanze della maggior parte degli elementi della Galassia derivano quasi esclusivamente dallo studio dei corpi del sistema solare. Il Sole però è tanto simile a molte altre stelle che si può ammettere che le quantità relative dei vari elementi nel sistema solare siano tipiche della grande maggioranza dei sistemi stellari. Le osservazioni del Sole rivelano che il 99,996 per cento della massa di un campione della sua materia consisterebbe di 15 elementi. L'elemento meno abbondante di questi 15, il nichel, sarebbe però ancora più comune degli altri 80 messi insieme.

Il modello della condensazione in equilibrio presuppone che non vi fossero sostanziali variazioni di temperatura nella nebulosa in corrispondenza dei pianeti che stavano accumulando materia solida sulla loro superficie. Forse il materiale si accumulò rapidamente in confronto alla velocità a cui la nebulosa stava raffreddando. O forse la composizione della materia solida era determinata dalla temperatura quando reagì definitivamente coi gas della nebulosa; dopo questo momento i gas potrebbero essersi dispersi e i solidi essersi lentamente aggregati sui pianeti ciascuno dei quali avrebbe avuto una composizione omogenea. In entrambi i casi l'equilibrio chimico tra gas e polvere avrebbe determinato la composizione sia dei gas sia dei prodotti di condensazione presenti in ogni momento.

Accanto ai disegni dei pianeti in dimensione relativa sono riportati i dati riguardanti ciascuno di essi e ottenuti sia dalle varie missioni spaziali sia dalle osservazioni da terra. Il segno negativo in corrispondenza del periodo di Venere indica che il pianeta ruota in direzione opposta a quella di rotazione di tutti gli altri pianeti.

GRADI KELVIN	MODELLO DELLA CONDENSAZIONE IN EQUILIBRIO	MODELLO DELL'ACCREZIONE NON OMOGENEA
1600	1. Condensazione di ossidi refrattari come ossido di calcio (CaO) e ossido di alluminio (Al_2O_3) come pure di ossido di titanio e ossidi delle terre rare	1. Condensazione di ossidi refrattari come ossido di calcio (CaO) e ossido di alluminio (AL_2O_3) come pure di ossido di titanio e ossidi delle terre rare
1300	2. Condensazione di lega metallica ferro-nichel	2. Condensazione di lega metallica ferro-nichel
1200	3. Condensazione del minerale enstatite ($MgSiO_3$)	3. Condensazione del minerale enstatite ($MgSiO_3$)
1000	4. Reazione del sodio (Na) con l'ossido di alluminio e silicati per dare feldspati e minerali associati, e deposito di potassio e altri metalli alcalini	4. Condensazione di ossido di sodio (Na_2O) e di altri ossidi di metalli alcalini a circa 800 gradi Kelvin
680	5. Reazione del gas acido solfidrico (H_2S) con ferro metallico per dare il minerale solfuro troilite (FeS)	
1200 - 490	6. Ossidazione progressiva del restante ferro metallico in ossido ferroso (FeO), che reagisce a sua volta con l'enstatite per dare olivina (Fe_2SiO_4 e Mg_2SiO_4)	
550	7. Combinazione del vapore acqueo (H_2O) con minerali che contengono calcio per dare tremolite	
425	8. Combinazione del vapore acqueo con l'olivina per dare serpentino	
175	9. Condensazione del vapore acqueo in ghiaccio	5. Condensazione del vapore acqueo in ghiaccio
150	10. Reazione del gas ammoniaca (NH_3) col ghiaccio d'acqua per dare l'idrato solido $NH_3 \cdot H_2O$	6. Condensazione del solfuro acido di ammonio (NH_4SH)
120	11. Reazione parziale del gas metano (CH_4) col ghiaccio d'acqua per dare l'idrato solido $CH_4 \cdot 7H_2O$	7. Condensazione del ghiaccio di ammoniaca (NH_3)
65	12. Condensazione dell'argo (A) e del rimanente gas metano in argo e metano solidi	8. Condensazione dell'argo (A) e metano solidi
20	13. Condensazione del neo (Ne) e dell'idrogeno che porta alla completa condensazione del 75 per cento della materia solare	9. Condensazione del neo (Ne) e dell'idrogeno che porta alla completa condensazione del 75 per cento della materia solare
~ 1	14. Condensazione dell'elio (He) in liquido	10. Condensazione dell'elio (He) in liquido

Le principali reazioni, che sarebbero avvenute nel corso della formazione del sistema solare secondo due diverse ipotesi, collegate alle varie temperature della nebulosa solare primordiale. Il modello della condensazione in equilibrio parte dall'ipotesi che la nebulosa si sia raffreddata molto lentamente durante la formazione dei pianeti. Il modello dell'accrezione non omogenea parte invece dall'ipotesi opposta: la nebulosa si raffreddò rapidamente rispetto alla velocità a cui si formavano i pianeti. Secondo l'autore, il modello della condensazione in equilibrio dà un quadro piú preciso dell'origine dei pianeti. L'area in colore indica due stadi di ciascuna ipotesi che probabilmente non si sono riscontrati; la temperatura infatti non è mai scesa a livelli cosí bassi.

Confronto dei due modelli

Esiste un secondo modello relativo alla formazione dei pianeti terrestri dalla nebulosa solare. Questo modello, l'ipotesi dell'accrezione non omogenea, è esattamente l'opposto del primo. In questo caso la nebulosa solare raffreddò rapidamente in confronto alla velocità alla quale i pianeti accumularono materiale solido. Siccome la composizione del materiale sarebbe stata determinata dalla sua temperatura, questo accumulo sul pianeta avrebbe dato origine a strati di diversa condensazione simili a quelli di una cipolla. È importante aggiungere che le due ipotesi sono casi estremi: la verità può trovarsi in qualsiasi punto intermedio. Siccome però non si hanno basi a priori per scegliere questa o quella, dovremmo comprendere le conseguenze di entrambi questi estremi.

I risultati delle due ipotesi sono completamente diversi (*si veda la figura a sinistra*). Se si parte con una temperatura di 2000 °K nella nebulosa solare, l'ipotesi della condensazione in equilibrio porterebbe prima alla condensazione di composti refrattari contenenti ossido di calcio (CaO), ossido di alluminio (Al_2O_3), ossidi delle terre rare e cosí via. Poi, quando la nebulosa si è raffreddata a circa 1500 gradi, si condenserebbe una lega di ferro-nichel simile a quella trovata nelle meteoriti. Questo stadio sarebbe seguito dalla condensazione di enstatite ($MgSiO_3$) e poi dalla formazione di vari minerali come feldspati attraverso la deposizione di sodio, potassio e altri metalli alcalini. Al diminuire della temperatura a 680 gradi il ferro metallico sarebbe corroso dal gas acido solfidrico (H_2S) per produrre il minerale troilite (FeS); il ferro rimanente verrebbe progressivamente ossidato per formare minerali come l'olivina che, combinandosi col vapor acqueo a temperatura ancora inferiore, darebbe serpentino. In fine, quando la temperatura della nebulosa scese sotto i 170 gradi, il vapor acqueo si sarebbe condensato in ghiaccio. Il ghiaccio avrebbe reagito poi col gas ammoniaca per fare un idrato di formula $NH_3 \cdot H_2O$. A 100 gradi il ghiaccio d'acqua si sarebbe combinato con parte del gas metano della nebulosa per dare un altro idrato: $CH_4 \cdot 7H_2O$. Alla fine l'argo e il rimanente gas metano sarebbero congelati a circa 60 gradi. Se la temperatura avesse continuato a scendere fino a 10 gradi si sarebbero condensati neo e idrogeno e, se la temperatura fosse scesa sotto un grado, persino l'elio. Non ci sono però prove che la temperatura entro la nebulosa solare sia mai scesa a livelli cosí bassi.

Come differisce questa sequenza di eventi dalla sequenza prevista dall'ipotesi dell'accrezione non omogenea? Una differenza importante consiste nel fatto che non si sarebbero formati quei minerali che possono essere generati solo dalle reazioni chimiche tra gas e minerali esistenti in precedenza. Questa limitazione esclude processi come la corrosione di ferro metallico da parte del vapor acqueo e da parte del gas acido solfidrico e la reazione di ghiaccio d'acqua con ammoniaca e metano per formare gli idrati $NH_3 \cdot H_2O$ e $CH_4 \cdot 7H_2O$. Cosí la sequenza di reazioni chimiche nei processi di accrezione non omogenea è molto piú semplice di quella che si ha nel processo della condensazione in equilibrio.

In primo luogo la condensazione di ossidi refrattari, della lega ferro-nichel e dell'enstatite si sarebbe svolta nello stesso modo in cui si sarebbe svolta nei processi della condensazione in equilibrio. Si sarebbe perciò condensato ossido di sodio (Na_2O); poi l'acqua, il solfuro acido di ammonio (NH_4SH) e l'ammoniaca sarebbero gelati nei rispettivi ghiacci. Alla fine sarebbero gelati metano e argo. Anche in questo caso la temperatura non sarebbe probabilmente mai scesa tanto da permettere la condensazione di neo, idrogeno e elio. È chiaro però che, per il materiale formatosi per accrezione non omogenea, la dipendenza della composizione e della densità dalla temperatura sarebbe completamente diversa da quella relativa al materiale formatosi per condensazione in equilibrio.

Esistono possibilità di verificare le previsioni dei due modelli nei confronti della realtà. Per esempio, disponiamo di moltissime informazioni circa la composizione e la struttura di un pianeta: la Terra. Quale delle due ipotesi è in grado di spiegare meglio ciò che conosciamo di questo pianeta? Siccome i due modelli prevedono la presenza di sostanze volatili nel materiale che formò i pianeti, qual è l'ipotesi che meglio si accorda con le nostre conoscenze sulle atmosfere planetarie? Possiamo fare delle distinzioni tra le due ipotesi sulla base della notevole quantità di informazioni che possediamo sulle condriti carbonacee, le meteoriti piú vecchie e primordiali?

Prove dalle meteoriti e dai pianeti

Consideriamo innanzitutto la prova fornita dalle condriti carbonacee. Questa prova è fortemente a favore del processo della condensazione in equilibrio. A differenza della crosta della Terra, le cóndriti sono composte di materia primordiale che non ha subito processi di riscaldamento, fusione, differenziazione. La mineralogia e le abbondanze relative degli elementi di questa materia preplanetaria sono a nostra disposizione per un'analisi diretta. Un minerale che è quasi universale nelle meteoriti è la troilite, o FeS. La troilite rappresenta una caratteristica importante della sequenza di eventi nella condensazione in equilibrio, ma è assente dalla sequenza nell'accrezione non omogenea. I minerali che vengono subito dopo, quanto a distribuzione, sono il pirosseno e l'olivina, contenenti per lo piú tra il 5 e il 20 per cento di ossido ferroso (FeO). L'ossido ferroso è parte essenziale della sequenza della condensazione in equilibrio.

Le condriti carbonacee sono ricche di materiali volatili, inclusa una grande quantità di acqua legata ad altre sostanze come il serpentino. È interessante il fatto che non sia possibile distinguere gli spettri di molti pianetini dagli spettri di laboratorio delle condriti carbonacee. Ciò sarebbe stato prevedibile sulla base del modello della condensazione in equilibrio. Infine le condriti carbonacee sono altamente omogenee. Nel materiale di queste meteoriti non vi è assolutamente alcuna tendenza a essere distribuito in molti minerali distinti allo stato puro; essi tendono invece a essere ben mescolati.

Anche i dati disponibili sulle atmosfere planetarie sono assai piú facilmente comprensibili sulla base del processo della condensazione in equilibrio che su quello dell'accrezione non omogenea. Il contenuto di elementi volatili nei pianeti formatisi per accrezione non omogenea sarebbe zero fino al momento in cui l'acqua si sarebbe condensata dalla nebulosa solare; il materiale condensato sarebbe perciò acqua per piú del 60 per cento! D'altra parte il modello della condensazione in equilibrio prevede la giusta quantità di acqua per la Terra, mentre lascerebbe Venere estremamente arida; prevede inoltre un contenuto di acqua per Marte sei volte superiore a quello della Terra.

Entrambi i modelli danno una esatta previsione della quantità di carbonio presente nei pianeti. Sembrerebbe che la fonte di ossigeno per produrre biossido di carbonio sui pianeti terrestri dovesse essere l'ossido ferroso. Il contenuto di ossido ferroso in Venere, nella Terra e in Marte previsto dal modello della condensazione in equilibrio è sufficiente in tutti e tre i casi per produrre la quantità osservata di biossido di carbonio. Nella sequenza prevista dal modello dell'accrezione non omogenea è però totalmente assente l'ossido ferroso.

Le proprietà chimiche e fisiche fondamentali della Terra sono tutte a favore del modello della condensazione in equilibrio. Non si può distinguere la composizione chimica della Terra nel suo insieme da quella delle condriti carbonacee ricche di ferro. Sfortunatamente, per gli studiosi di questi problemi però, circa il 99 per cento della massa della Terra non è composto dal materiale della crosta, ma dal materiale del mantello e del nucleo e nessuno dei due è analizzabile direttamente. Ci si deve accontentare di dedurre la composizione di questi materiali da osservazioni indirette di alcune proprietà come la densità e la velocità delle onde sismiche.

Si è d'accordo in generale sul fatto che il mantello terrestre consista di composti contenenti ossido di magnesio (MgO), biossido di silicio (SiO_2) e ossido ferroso. Sembra probabile che tutto il mantello sia costituito per il 10 per cento di ossido ferroso. Il mantello si divide in due strati distinti: lo strato superiore formato soprattutto da silicati di ferro e magnesio e quello inferiore, in cui la pressione è piú alta, composto in larga misura di una miscela di ossidi di ferro e di magnesio.

Anche il nucleo si divide in due strati. Il nucleo interno, che comprende solo l'uno per cento della massa della Terra, ha una densità che corrisponde alla densità di una lega solida di ferro-nichel sottoposta a una pressione enorme. Il nucleo esterno è liquido e ha una densità apprezzabilmente minore di quella del ferro liquido puro alla stessa pressione e, quindi, vi deve essere miscelato al ferro un elemento piú leggero. Si pensa che questo componente piú leggero possa essere o silicio o zolfo. È già una notevole coincidenza che il modello della condensazione in equilibrio fornisca l'abbondanza corretta dello zolfo per spiegare la densità osservata del nucleo esterno. Inoltre il modello della condensazione in equilibrio spiega perché la Terra abbia una densità leggermente superiore a quella di Venere ricorrendo all'ipotesi che la Terra contenga zolfo nel nucleo, e Venere no.

Non sappiamo ancora se Venere abbia realmente meno zolfo rispetto alla Terra, ma è certamente vero che l'atmosfera di Venere non contiene composti di zolfo rilevabili. Se Venere emette gas vulcanici, come fa la Terra, dovrebbe contenere nella sua atmosfera una quantità facilmente rilevabile di solfuro di carbonile (COS). Il solfuro di carbonile non è però stato trovato

neanche con metodi spettroscopici assai sensibili.

Struttura interna dei pianeti

Sulla base di tutte le prove derivate dalle meteoriti e dai pianeti accettiamo il modello della condensazione in equilibrio come quello che dà la migliore approssimazione della differenza primordiale nella composizione del sistema solare. Esaminiamo ora brevemente la natura dei processi fisici e chimici che avvengono all'interno dei pianeti e dei satelliti per valutare quali tipi di struttura interna possano essersi formati nei corpi che hanno loro dato origine.

La domanda piú facile da porsi è come l'attuale struttura interna di oggetti differenziati come la Terra dipenda dalla temperatura a cui si sono formati. (Corpi piccoli come i meteoroidi e i pianetini, che non sono mai stati abbastanza caldi da differenziarsi, resterebbero naturalmente miscele omogenee dei loro minerali costituenti.) Per rispondere a questa domanda dovremmo passare in rassegna le fasi del processo della condensazione in equilibrio, quando la temperatura nella nebulosa solare scese dai 2000 gradi kelvin.

I primi minerali che si sono formati, gli ossidi refrattari, avrebbero dato origine a una classe di protopianeti con un'alta concentrazione di calcio, alluminio, titanio e terre rare. I protopianeti sarebbero stati ricchi anche di uranio e di torio e quindi si sarebbero scaldati rapidamente per il decadimento radioattivo di questi elementi nel loro interno. Cosí, nonostante il fatto che i minerali ossidi refrattari abbiano punti di fusione alti, la temperatura entro i protopianeti sarebbe stata abbastanza alta per fonderli facilmente. Attualmente l'interno di un corpo grande di questo tipo sarebbe per la gran parte omogeneo. I solfuri, il ferro e gli ossidi di ferro sarebbero totalmente assenti. È stato recentemente suggerito da Don L. Anderson dell'Istituto di tecnologia della California che la Luna sia un membro di questa classe di composizione. La Luna avrebbe però dovuto essere stata notevolmente contaminata da solfuri e ossidi di ferro per giustificare la composizione che si osserva.

La seconda classe di composizione sarebbe stata composta di protopianeti formatisi subito dopo la condensazione dalla nebulosa solare della lega metallica ferro-nichel. Dopo l'avvenuta differenziazione, questi corpi avrebbero avuto una sottile crosta di minerali ossidi poggiante direttamente sul massiccio nucleo metallico. Attualmente, nel sistema solare non sono presenti corpi di questo tipo.

I protopianeti della terza classe si sarebbero formati dopo che la temperatura della nebulosa era scesa abbastanza da permettere la condensazione della enstatite. A questo punto la composizione di un pianeta in via di formazione sarebbe stata dominata dai silicati. Una crosta di ossidi avrebbe formato uno straterello sopra un mantello ricco di pirosseni contenente una quantità trascurabile (meno dello 0,01 per cento) di ossido ferroso. Il pianeta Mercurio è un membro di questa classe.

I protopianeti della quarta classe sarebbero stati diversi da quelli della terza in quanto la loro crosta avrebbe conservato i metalli alcalini, col risultato di una composizione chimica simile a quella della crosta terrestre eccetto per il contenuto molto piú basso di elementi volatili. Un corpo di questo genere avrebbe avuto un mantello contenente, forse, lo 0,5 per cento di ossido ferroso. I corpi di questa classe avrebbero il corrispondente in Venere.

I protopianeti della quinta classe avrebbero conservato lo zolfo come componente della troilite. Liquefacendosi, la troilite sarebbe sprofondata formando un nucleo esterno denso. La maggior parte del nucleo sarebbe stata composta da questo solfuro fuso invece che da metallo solido. Il mantello, che sarebbe stato composto di olivina e di pirosseno, avrebbe contenuto almeno il due per cento di ossido ferroso. Questa classe non è rappresentata nel sistema solare attuale salvo forse nelle meteoriti.

La sesta classe principale sarebbe stata composta dai protopianeti formatisi quando la temperatura nella nebulosa solare era diminuita in modo sufficiente (a circa 600 gradi kelvin) per conservare l'acqua legata nella struttura cristallina di minerali come la tremolite. Fino allo 0,3 per cento la massa di questi corpi potrebbe essere stata acqua. La Terra, con un nucleo ricco di zolfo, con un mantello contenente circa il 10 per cento di ossido ferroso e col suo contenuto di acqua di circa lo 0,05 per cento, appartiene a questa classe.

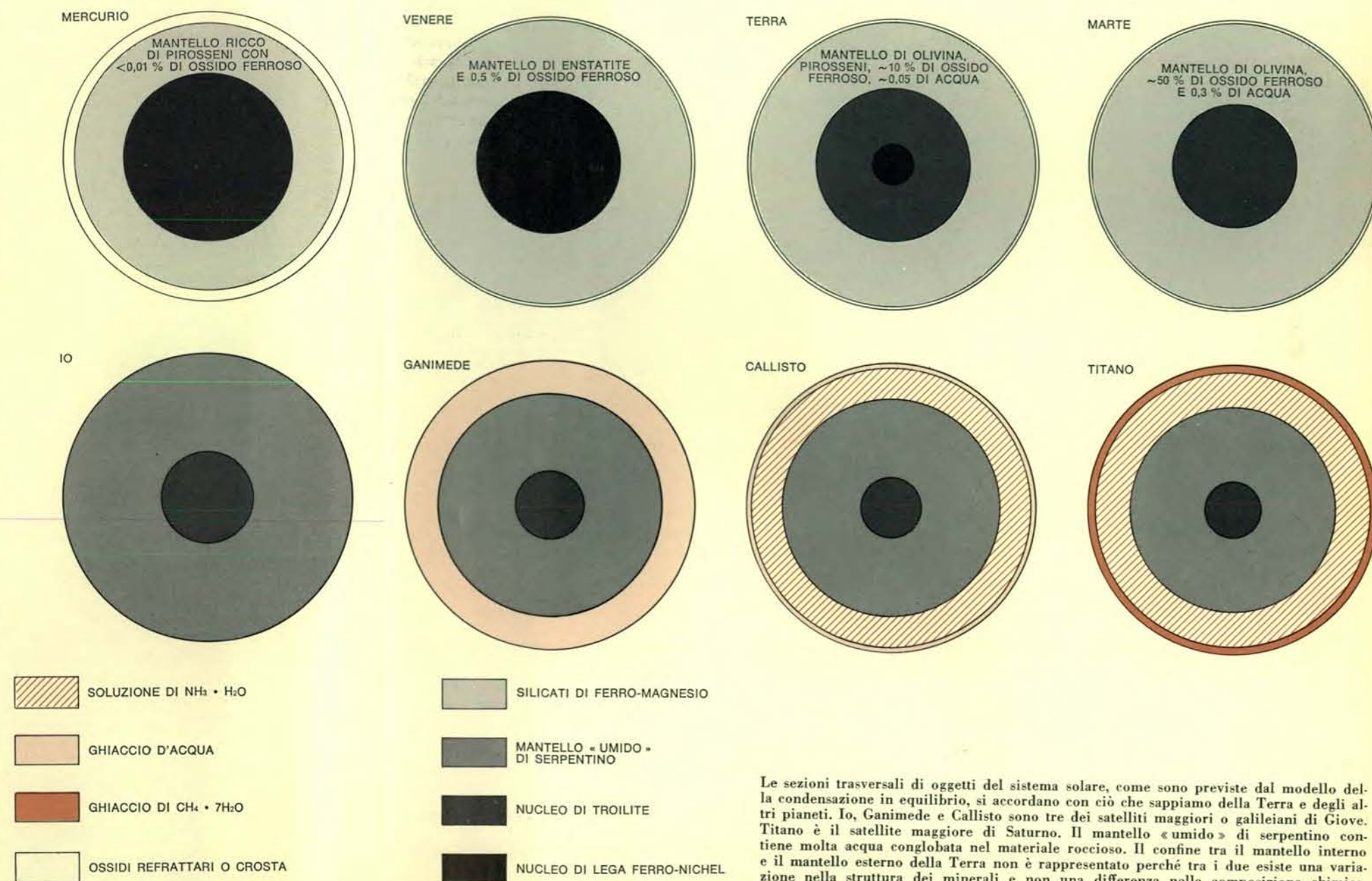
Marte e oltre

La settima classe principale sarebbe stata composta di protopianeti formatisi dopo che il ferro metallico era stato completamente ossidato. Un pianeta di questo genere avrebbe avuto un contenuto di circa lo 0,3 per cento di acqua. Il suo mantello di olivina sarebbe stato ricchissimo di ossido ferroso, probabilmente per un ammontare di circa il 50 per cento. Il nucleo sarebbe stato di troilite sia solida che liquida priva di ferro metallico. L'oggetto che ne sarebbe risultato assomiglierebbe molto a Marte.

I protopianeti dell'ottava classe, che si sarebbero formati a una temperatura abbastanza bassa perché l'olivina del mantello si combinasse col radicale idrossile della nebulosa solare per formare un mantello «umido» di serpentino, avrebbe avuto un contenuto di acqua spaventosamente alto: circa il 14 per cento della massa. La differenza di densità tra i minerali della crosta e quelli del mantello sarebbe essenzialmente scomparsa, ma è possibile che sia potuto esistere un nucleo di solfuro ferroso. Un corpo omogeneo originario con questa composizione sarebbe stato sufficientemente differenziato da formare un nucleo anche a questa temperatura relativamente bassa di 400 gradi, perché il punto di fusione dei silicati con alto contenuto di acqua è abbastanza basso. I pianetini appartengono probabilmente a questa ottava classe. I calcoli basati sui punti di fusione noti dei silicati dovrebbero dirci se questi piccoli oggetti potrebbero essersi riscaldati sufficientemente per il decadimento di elementi radioattivi nel loro interno perché i minerali si differenziassero nel loro debole campo gravitazionale. Se i pianetini fossero infatti simili alle condriti carbonacee, allora forse le ordinarie condriti piú secche e meno ossidate ebbero origine tra le orbite della Terra e di Marte invece che nella fascia degli asteroidi.

I protopianeti della nona classe, formatisi a temperature inferiori a quella di condensazione dell'acqua, avrebbero contenuto acqua ghiacciata. In un corpo di questo tipo il contenuto d'acqua, sia come ghiaccio sia come acqua legata nei minerali, sarebbe stato circa il 60 per cento in peso. La densità risultante sarebbe stata perciò assai bassa: solo circa 1,7 grammi per centimetro cubo e meno di un terzo della densità media della Terra. Il satellite maggiore di Giove, Ganimede, potrebbe appartenere a questa classe.

I protopianeti della decima classe avrebbero conservato l'ammoniaca come idrato solido $NH_3 \cdot H_2O$. Una miscela di questo idrato con ghiaccio comincia a fondere alla bassa temperatura di 173 gradi kelvin, che è solo di pochi gradi superiore alla temperatura diurna superficiale del secondo satellite in ordine di grandezza di Giove, Callisto. Se un protopianeta di questa classe fosse stato sufficientemente scaldato dal decadimento radioattivo in modo da differenziarsi, avrebbe avuto un mantello profondo di una soluzione liquida di acqua-ammoniaca ricoperto da una sottile crosta di ghiaccio d'acqua. Durante la differenziazione si sarebbe formato, come sedimento, un nucleo di solfuri, ossidi e silicati idrati (contenenti acqua). Il risultato globale sarebbe stata una palla stratificata di fango e di argilla.

I protopianeti dell'undicesima classe avrebbero avuto circa il 4 per cento in peso di metano. Il calore generato dal decadimento radioattivo avrebbe fuso l'idrato solido $CH_4 \cdot 7H_2O$ liberando gas metano per formare un'atmosfera. La differenziazione avrebbe dato ori-

Le sezioni trasversali di oggetti del sistema solare, come sono previste dal modello della condensazione in equilibrio, si accordano con ciò che sappiamo della Terra e degli altri pianeti. Io, Ganimede e Callisto sono tre dei satelliti maggiori o galileiani di Giove. Titano è il satellite maggiore di Saturno. Il mantello «umido» di serpentino contiene molta acqua conglobata nel materiale roccioso. Il confine tra il mantello interno e il mantello esterno della Terra non è rappresentato perché tra i due esiste una variazione nella struttura dei minerali e non una differenza nella composizione chimica.

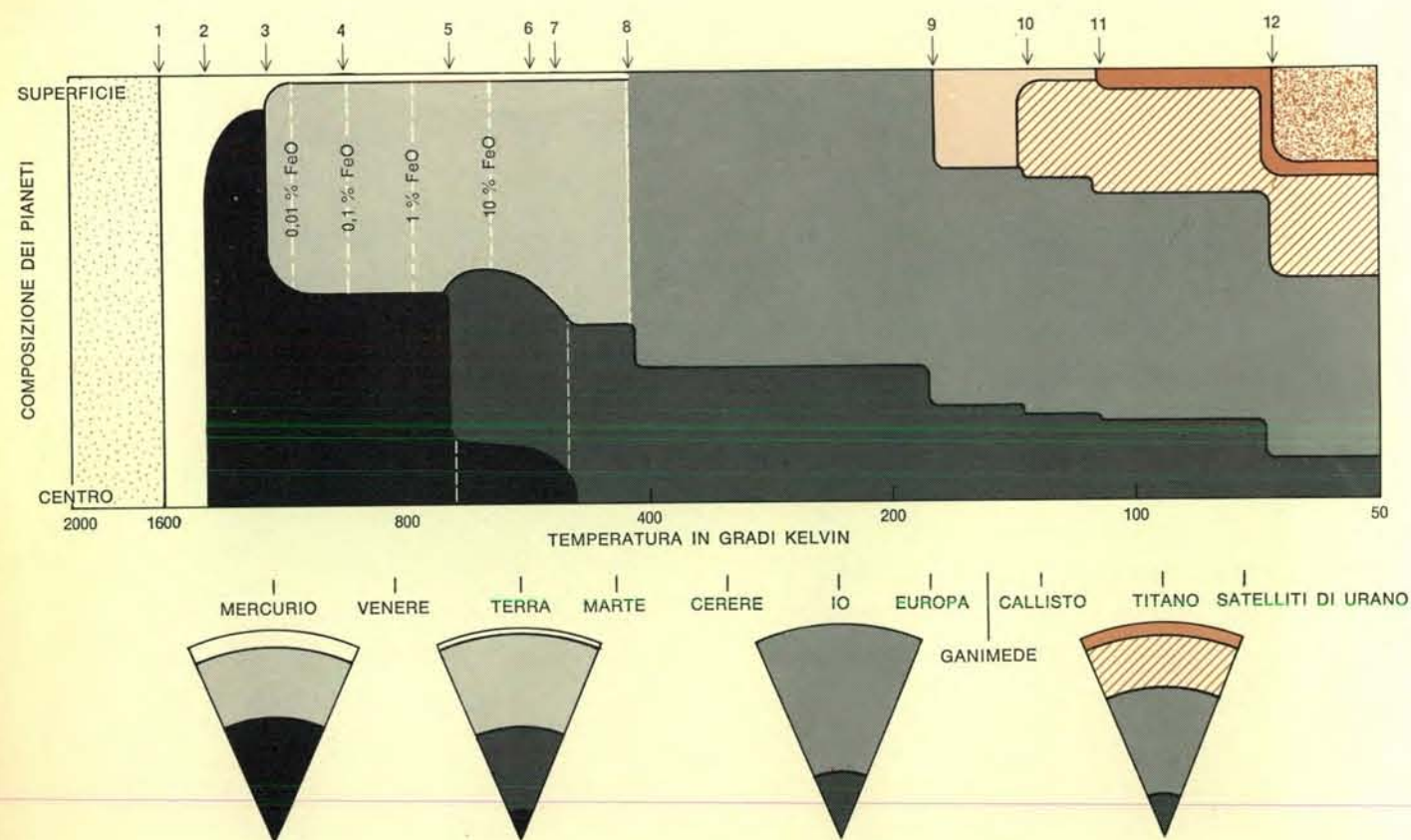
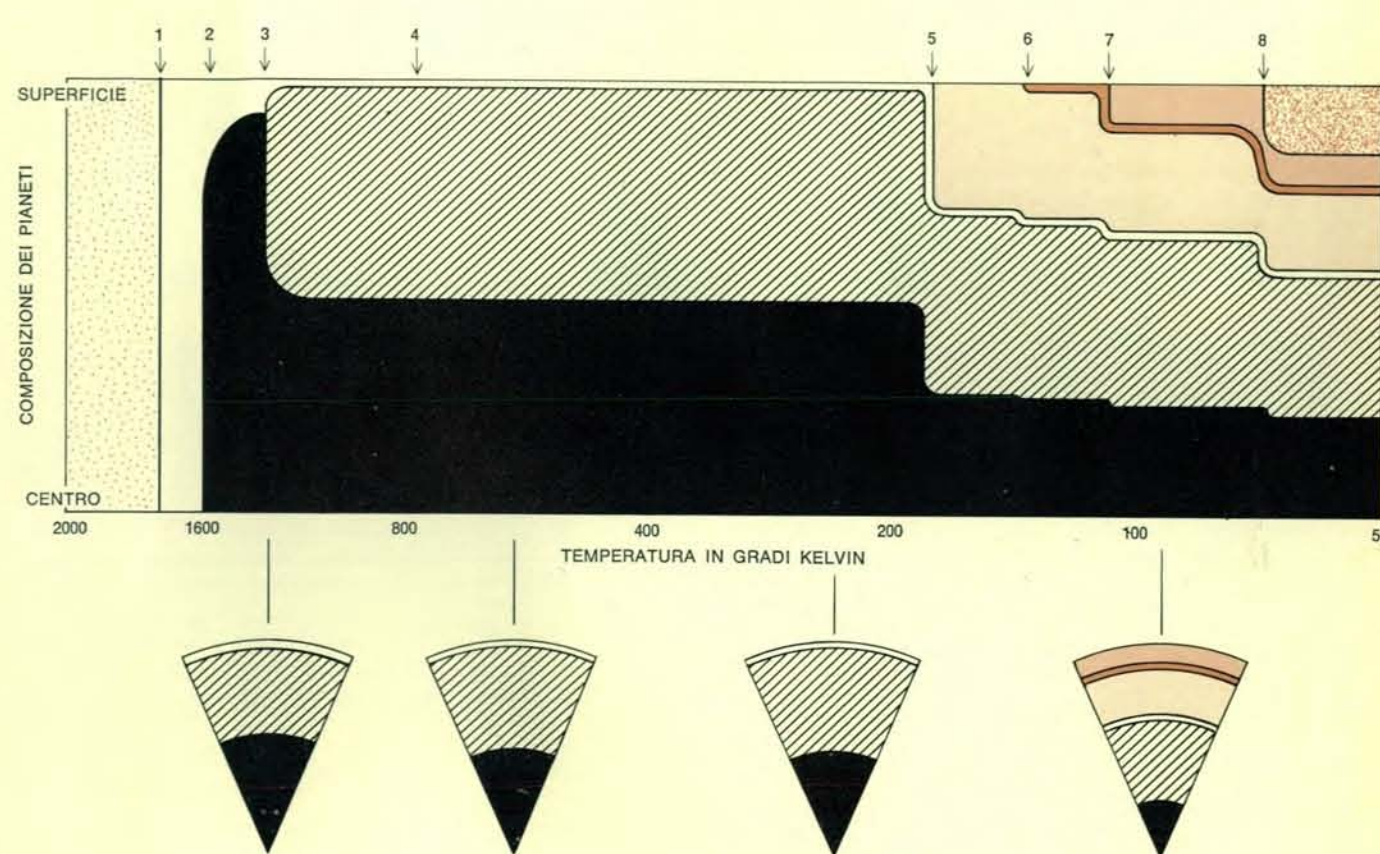
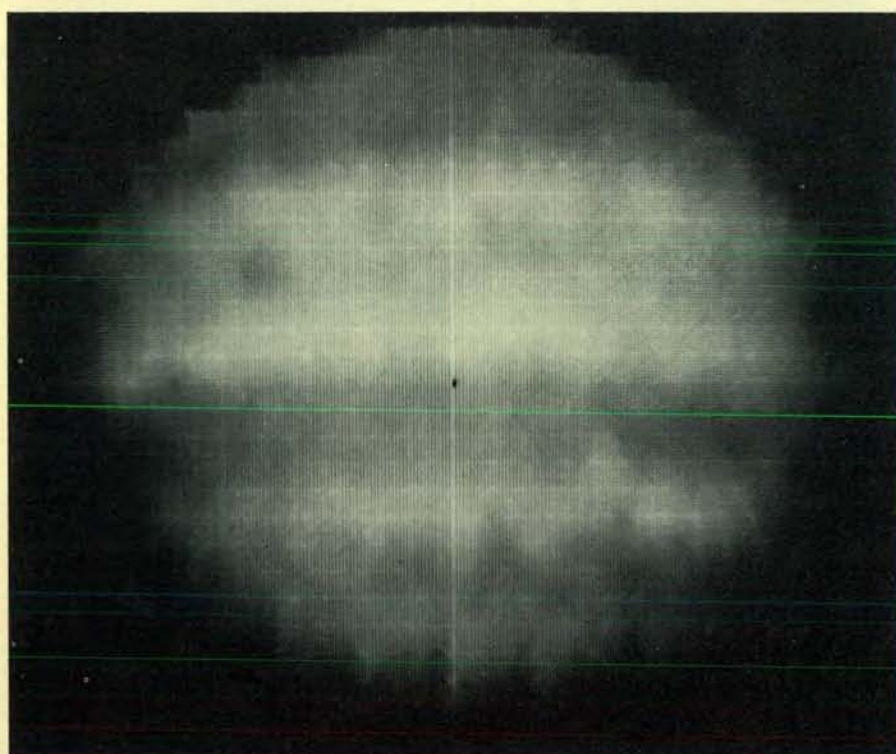
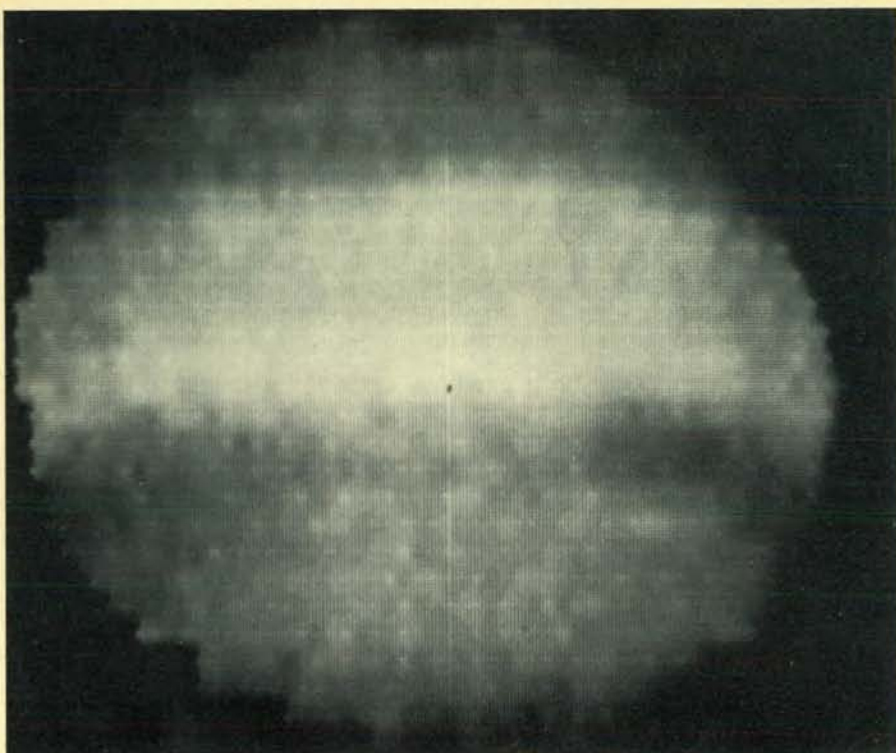
gine a una crosta solida di idrato. Un primo candidato come membro di questa classe è il satellite maggiore di Saturno, Titano, che ha una densa atmosfera ricca di metano.

I protopianeti della dodicesima e ultima classe avrebbero trattenuto metano solido quando la temperatura della nebulosa solare scese a circa 50 gradi kelvin. Siccome il carbonio ha, nel sistema solare, circa la stessa abbondanza dell'ossigeno e siccome la densità del metano solido o liquido è bassissima (circa 0,6 grammi per centimetro cubo), la densità di un tale protopianeta sarebbe stata solo un grammo per centimetro cubo. Si dice che molti dei satelliti minori di Saturno abbiano densità vicine a un grammo per centimetro cubo, ma il margine di errore di queste misure è almeno di un fattore due. È perfettamente possibile che alcuni satelliti di Urano e di Nettuno appartengano a questa classe, ma le loro masse e i loro raggi dovono essere determinati da missioni spaziali interplanetarie prima di poter confermare con una qualche certezza questa ipotesi.

Ho citato i satelliti di Giove, Saturno, Urano e Nettuno, ma non i pianeti stessi. Vi sono ragioni atte a spiegare l'assenza dei pianeti gioviani dalle classi di composizione. Come abbiamo visto queste classi si basano sulla temperatura a cui il materiale protoplanetario condensò dalla nebulosa solare secondo il modello della condensazione in equilibrio. Ora, siccome la composizione di Giove e di Saturno è essenzialmente uguale alla composizione solare, essi potrebbero essersi formati quasi ovunque nella nebulosa solare senza che esistesse una stretta dipendenza tra quella regione e la temperatura. Tutto quello che sarebbe richiesto è un nucleo protoplanetario iniziale il cui campo gravitazionale fosse abbastanza intenso da accumulare materiale solare non differenziato. Urano, Nettuno e Plutone sono stati esclusi semplicemente perché non si sa ancora abbastanza sulla loro composizione.

Risultati importanti

Un risultato importante della sequenza di reazioni che abbiamo appena discusso è che la densità dei pianeti terrestri, come la loro chimica, si può spiegare come conseguenza diretta della variazione di temperatura a cui si sono formati secondo la distanza dal Sole. Si è spesso affermato che, per analogia con l'ipotesi sulla formazione delle meteoriti, le densità diverse dei pianeti terrestri sono dovute a un processo fisico che ha frazionato il materiale in modo tale che i metalli si accumulassero preferenzialmente in un caso (Mercurio) e i silicati si accumulassero preferenzialmente negli altri casi. Questo accumulo preferenziale, come si osserva nelle meteoriti, può essere stato il risultato di differenze nelle proprietà magnetiche o elettriche del materiale o di come la materia accumulata ha aderito bene alla superficie del protopianeta. Siccome la sequenza di reazioni del modello della condensazione in equilibrio porta automaticamente all'alta densità di Mercurio, siamo portati a negare che un qualsiasi processo di frammentazione abbia interessato i pianeti. I pianeti sono cosí massicci che devono aver accumulato materia per gravitazione, un processo totalmente non selettivo. La temperatura da sola determinava la composizione del materiale che si accumulava. D'altra parte i corpi che hanno dato origine alle meteoriti e che avrebbero avuto una massa fino a 10^{-12} volte inferiore a quella di un pianeta sarebbero stati dotati solo di una piccola attrazione gravitazionale. Accumularono probabilmente il loro materiale quasi interamente attraverso forze selettive tra particelle come il magnetismo.

In conclusione, il modello della condensazione in equilibrio mostra in modo molto semplice come la densità e il contenuto volatile del materiale solido nel sistema solare siano legati alla temperatura a cui si è formato il materiale. Il modello suggerisce una notevole varietà di previsioni quantitative che possono essere verificate in laboratorio, con osservazioni astronomiche e con sonde planetarie. Entro una singola sequenza chimica possiamo ora legare tra loro la composizione della nebulosa solare, la struttura interna dei pianeti e il loro contenuto di composti volatili.

Il modello della condensazione in equilibrio per la formazione del sistema solare è messo a confronto con il modello dell'accrezione non omogenea (*si veda la figura nella pagina a fronte*). Il bordo a sinistra del diagramma, a 2000 gradi kelvin, rappresenta l'interno della nebulosa solare; il bordo a destra, a 50 gradi kelvin, rappresenta una distanza pari a circa 30 volte la distanza della Terra dal centro della nebulosa. I numeri in alto indicano le posizioni delle 12 reazioni principali dell'ipotesi della condensazione in equilibrio. In basso sono indicate le temperature a cui si sarebbero formati alcuni oggetti del sistema solare. L'asse verticale rappresenta le previsioni del modello per la composizione interna di oggetti dal centro (*in basso*) alla superficie (*in alto*) indipendentemente dalla particolare distanza in chilometri. Gli spicchi in basso mostrano le sezioni trasversali di quattro oggetti del sistema solare (Mercurio, Terra, Io e Titano) che si formarono a diverse temperature. L'area della lega ferro-nichel rappresenta il nucleo interno della Terra; l'area di troilite è il nucleo fuso piú esterno.

Il modello dell'accrezione non omogenea per la formazione del sistema solare prevede composizioni dei pianeti completamente diverse dato che essi si formarono a varie temperature. Questo diagramma è analogo a quello della pagina a fronte. Le otto reazioni principali del modello dell'accrezione non omogenea sono raffigurate in alto. Gli spicchi in basso rappresentano le sezioni trasversali di ipotetici pianeti che si sarebbero formati secondo il modello dell'accrezione non omogenea alla stessa temperatura a cui si formarono Mercurio, Terra, Io e Titano secondo il modello della condensazione in equilibrio. L'ipotetico Mercurio avrebbe avuto la stessa struttura. La Terra però avrebbe avuto un mantello di minerale enstatite e un nucleo solido di lega ferro-nichel, previsione che non è confermata dagli studi sulle onde sismiche. L'ipotetico Io avrebbe avuto la stessa composizione di Venere, della Terra e degli asteroidi anche se tutti si formarono a temperature e distanze dal Sole profondamente diverse. Titano sarebbe stato costituito di una mezza dozzina di strati fatti di materiali diversi.

Mappe termiche di Giove a due diverse lunghezze d'onda fatte da *Pioneer 10* con l'esperimento del radiometro infrarosso dell'Istituto di tecnologia della California diretto da Guido Münch. Il radiometro è un telescopio riflettore con uno specchio di 7,5 centimetri montato rigidamente sul veicolo spaziale. La rotazione di *Pioneer 10* e il suo moto sulla traiettoria oltre Giove permisero al radiometro di analizzare il pianeta e di ricostruirne le immagini. La fotografia in alto è stata fatta a una lunghezza d'onda di 20 micron ed è una misura del calore emesso a una distanza di 40 chilometri sotto lo strato superiore dell'atmosfera di Giove; la fotografia in basso è stata fatta a una lunghezza d'onda di 40 micron ed è una misura del calore emesso a una distanza di 50 chilometri sotto il limite superiore dell'atmosfera. Entrambe furono fatte il 3 dicembre 1973, durante il massimo avvicinamento del veicolo spaziale al pianeta. Le aree piú calde (piú chiare) hanno una temperatura di otto gradi kelvin superiore a quella delle aree piú fredde (piú scure). Il pianeta è piú caldo vicino all'equatore. Si pensa che il pianeta abbia una temperatura media di circa 128 gradi kelvin; esso irradia piú del doppio della quantità di energia termica che riceve dal Sole, confermando le osservazioni fatte da terra che indicano che il pianeta stesso genera calore.

Il risultato forse piú sorprendente di questo approccio è però il quadro di base di un'intima correlazione tra tutti gli oggetti del sistema solare. Osserviamo che vi sono forti indizi che i pianeti si siano formati da un mezzo di composizione uniforme. Osserviamo che gli stessi processi qualitativi di condensazione e accrezione operano nello stesso modo generale in tutto il sistema. La presenza o l'assenza di gas di zolfo nell'atmosfera di Venere, la densità del nucleo di Marte, la conduttività del mantello di Callisto, la pressione atmosferica alla superficie di Titano, il raggio e la massa dei satelliti di Urano, tutti influenzano in modo tangibile la comprensione dell'origine, della composizione e della struttura del nostro pianeta. In questo momento della storia della geologia abbiamo constatato quanto sia importante lo studio di tutti gli elementi chimici per comprendere l'origine di ciascuno di essi cosí come siamo stati obbligati a studiare tutta la tettonica per riuscire a comprendere l'origine di una singola montagna.

Dobbiamo ora essere preparati a imparare dallo studio comparato dei pianeti quali processi esattamente influenzano i pianeti in generale e come i pianeti evolvono. Da questo tipo di ricerca possiamo apprendere per la prima volta come la Terra abbia avuto origine, si sia evoluta e sia giunta alla sua struttura attuale.

Un altro vantaggio di larga portata che deriva da tali studi comparati sarà quello di comprendere con quale frequenza si possano essere formati pianeti intorno ad altre stelle e quale possa essere la loro composizione. Siccome la composizione della maggior parte delle stelle è identica a quella del Sole e siccome le teorie attuali sull'origine del sistema solare suggeriscono che la formazione dei pianeti sia un normale sottoprodotto della formazione delle stelle, non sarebbe sorprendente scoprire che pianeti simili a quelli del sistema solare sono comuni nella nostra Galassia e in altre galassie a spirale. È interessante considerare se esistano o meno in altri sistemi planetari pianeti che appartengano a qualche classe di composizione che non si osserva nel sistema solare.

L'immediato futuro offrirà numerose opportunità per verificare direttamente queste previsioni. Un vasto programma di esplorazione dei pianeti è il solo modo di sfruttare al massimo le opportunità che ci sono offerte: è la logica prosecuzione dei nostri tentativi diretti alla comprensione del pianeta su cui viviamo.

Polimeri inorganici

I polimeri inorganici, ancora in fase di studio, presentano caratteristiche insolite e diverse da quelle dei polimeri organici impiegati nell'industria

di Harry R. Allcock

I prodotti della chimica dei polimeri sono tutti intorno a noi: vestiti confezionati con fibre sintetiche, recipienti di polistirene, bottiglie di polietilene, ingranaggi di nylon, pentole da cucina rivestite in Teflon, sacchi di plastica, vernici a base di polimeri, colle epossidiche, gomma sintetica, schiume poliuretaniche – la lista è quasi senza fine. I vent'anni trascorsi hanno visto una rivoluzione tecnologica nella quale i materiali tradizionali, quali cotone, legno, vetro e metalli, sono stati sostituiti dai polimeri sintetici.

Quasi tutte le materie plastiche e gli elastomeri impiegati oggi sono polimeri organici. Essi sono costituiti da molecole molto grandi, contenenti atomi di carbonio disposti a formare catene lineari o ramificate. Per esempio le molecole di polietilene sono costituite da lunghe catene di atomi di carbonio a ciascuno dei quali sono legati due atomi di idrogeno. Il polistirene ha una struttura simile ad eccezione del fatto che un anello fenilico (C_6H_5) sostituisce un atomo di idrogeno ogni due atomi di carbonio. Catene lineari sono presenti anche nel polipropilene, nel Nylon 66, nel polimetilmetacrilato (Lucite o Plexiglas) e nel polietilentereftalato (Dacron e Mylar). Polimeri sintetici come quelli citati possono essere impiegati per fabbricare fibre tenaci, forme solide complicate, film resistenti ed elastomeri simili alla gomma.

Rimane però il fatto che molti polimeri organici presentano gravi deficienze. Pochi possono essere riscaldati per periodi prolungati sopra i 150 gradi centigradi senza fondere o decomporsi. Quando avviene la decomposizione, essa è generalmente il risultato della reazione degli atomi di carbonio con l'ossigeno atmosferico. Molti polimeri organici si sciolgono o rigonfiano in solventi organici a caldo o in oli minerali e per questa ragione non possono essere usati per guarnizioni o componenti dei motori d'automobile e d'aeroplano.

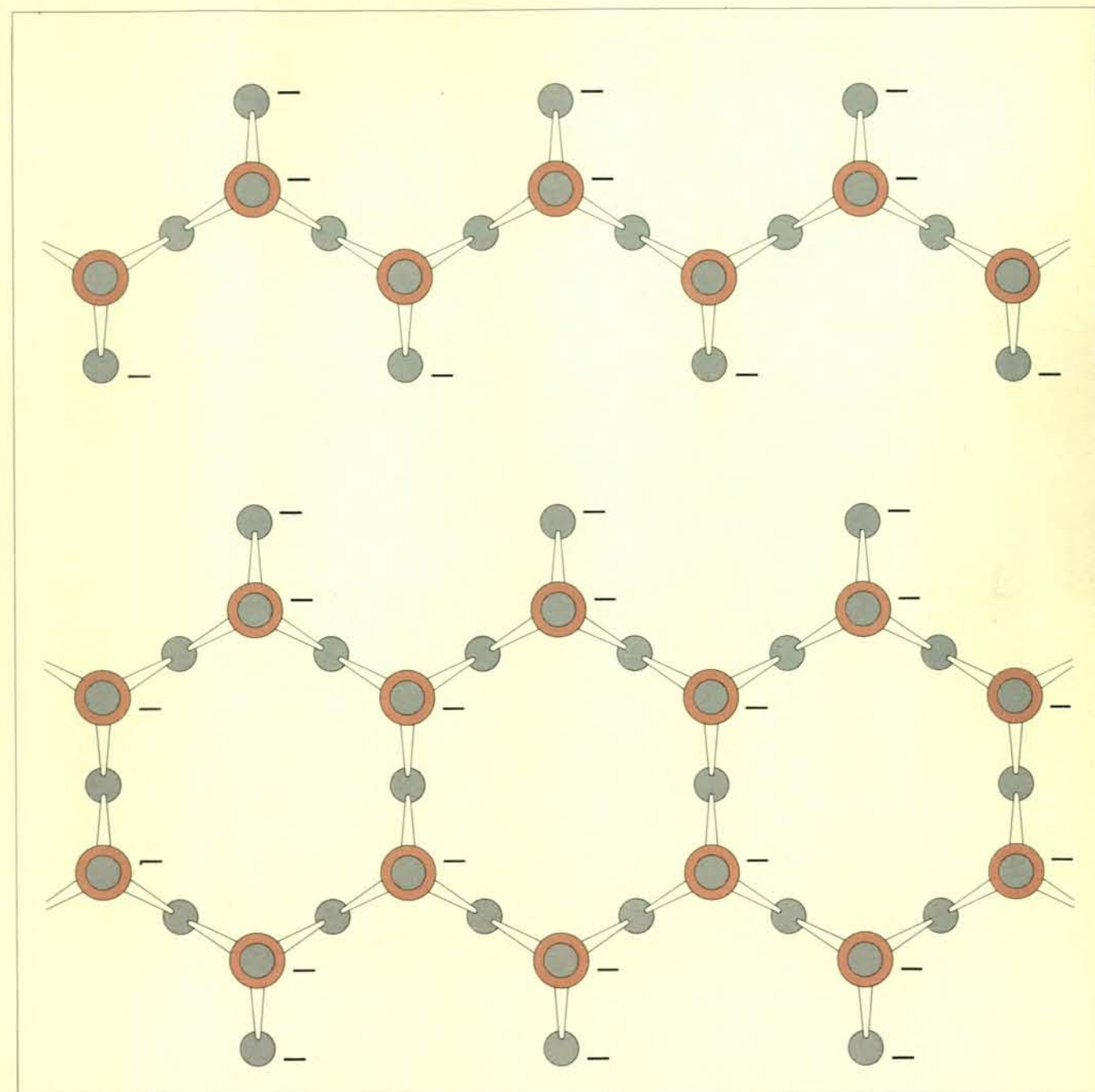
Inoltre, pochi polimeri organici rimangono flessibili o gommosi in un intervallo di temperatura abbastanza ampio per poter essere impiegati sia alle basse sia alle alte temperature. Questa è la ragione per cui i pneumatici delle automobili che contengono tortiglie di nylon sono maggiormente rigidi nelle mattine fredde, e i tubi flessibili di plastica per giardino in inverno non possono essere arrotolati. Nell'Artico o negli aerei che volano ad alta quota, questa durezza e fragilità dei polimeri può creare seri pericoli. Vi sono altre applicazioni per le quali non sono stati trovati tuttora polimeri organici adatti: nei voli spaziali, dove è essenziale una resistenza prolungata alla radiazione ultravioletta; nell'industria tessile dove sono richiesti tessuti antifiamma, e in medicina per la produzione di organi artificiali che non favoriscano la coagulazione del sangue o provochino altri effetti indesiderabili. Molti studiosi di polimeri sono convinti che la soluzione di questi problemi richieda lo sviluppo di sistemi polimerici radicalmente differenti. Questo spiega perché l'attenzione è sempre più focalizzata sui polimeri inorganici.

La maggior parte dei polimeri sintetici a catena lunga mostra una caratteristica successione di modificazioni quando viene riscaldata. Tutti i polimeri lineari sono vetrosi alle basse temperature. Quando vengono riscaldati si raggiunge una certa temperatura alla quale essi si trasformano da vetrosi a gommosi. Questo cambiamento è conosciuto come temperatura di transizione vetrosa. A temperature ancora più elevate il polimero può rammollire lentamente ed eventualmente liquefarsi. Se il materiale è un polimero cristallino, come avviene se segmenti di catene adiacenti sono « impacchettati » insieme in una struttura regolare, esso fonderà in un liquido viscoso a una temperatura ben definita. La formazione di estesi legami incrociati tra le catene porta a un polimero reticolato, un materiale rigido con punto di fusione elevato (*si veda l'illustrazione a pagina 48*).

Le differenze rilevabili tra i polimeri sono attribuibili principalmente a differenti temperature di transizione vetrosa, poiché variano il grado di cristallinità e le caratteristiche specifiche di fusione. Queste differenze sono dovute in larga misura alla presenza di gruppi sostituenti diversi legati alla catena principale, oppure alla presenza, nella catena, di elementi diversi dal carbonio. Così, variando la struttura dei gruppi laterali o introducendo atomi di ossigeno o di azoto in una catena di atomi di carbonio, è stato possibile ottenere una stupefacente varietà di polimeri con un'enorme gamma di proprietà diverse.

Fondamentalmente i polimeri inorganici dovrebbero mostrare le stesse transizioni termiche dei polimeri organici. Per esempio, dovrebbero avere aspetto vetroso a basse temperature, diventare gommosi al di sopra della temperatura di transizione vetrosa e infine diventare viscosi e fondere. Inoltre l'estesa reticolazione delle catene dovrebbe portare a matrici rigide che hanno un alto punto di fusione. Nello stesso tempo i polimeri inorganici dovrebbero presentare una più considerevole resistenza alla degradazione ossidativa rispetto ai polimeri organici.

Da lungo tempo vengono usati alcuni sistemi polimerici inorganici molto importanti. Il vetro è un polimero inorganico costituito da anelli e catene di unità silicato ripetentesi. È noto che le fibre di vetro possono essere usate per la fabbricazione di articoli tessili. Catene similari di unità silicato si trovano nei minerali di pirossene. Nei minerali di anfibolo si trovano polimeri reticolati o a doppia catena come quelli che formano gli asbesti (*si veda l'illustrazione in questa pagina*). In entrambe queste strutture le cariche negative degli atomi di ossigeno che sono situati al di fuori della catena centrale sono neutralizzate da ioni metallici caricati positivamente, quali gli ioni sodio, magnesio, calcio. Queste strutture laterali polari tendono a legare insieme catene adiacenti per formare una matrice reticolata. Così, invece di presentare proprietà gommose ed elastomeriche o proprietà di flessibilità e capacità di formazione di film, i polimeri hanno un elevato punto di fusione. Quando sono presenti legami incrociati covalenti, come nei vetri silicei, un sistema tridimensionale di legami covalenti silicio-ossigeno, genera una matrice ad altissimo punto di fusione.

Sono inoltre conosciuti polimeri fosfatici che hanno uno scheletro in cui atomi di fosforo si alternano ad atomi di ossigeno e questi formalmente assomigliano ai silicati lineari. Per esempio, alti polimeri cristallini di fosfati alcalini sono conosciuti da parecchi anni quali « sali di Maddrell » e « sali di Kurrol ». Inoltre la presenza di gruppi laterali polari abbassa la mobilità di torsione della catena, con conse-

Catene e anelli inorganici presenti in certe classi di minerali, offrono modelli utili allo studioso dei polimeri. Minerali di pirossene (*sopra*) sono polimeri lineari in cui atomi di ossigeno e atomi di silicio (*in colore*) si alternano nello scheletro della catena. Ciascun atomo di silicio è inoltre legato a due atomi di ossigeno che si proiettano al di fuori dello scheletro. Questi atomi di ossigeno portano una carica negativa che viene neutralizzata dalla carica positiva degli ioni metallici quali sodio (*non indicati*). La disposizione delle cariche lega le catene adiacenti l'una all'altra per formare una matrice ad alto grado di reticolazione. I minerali di anfibolo (*sotto*) presentano una struttura a doppia catena reticolata secondo tale sistema.

guente eliminazione di flessibilità a bassa temperatura. Un altro problema incontrato con alcuni di questi polifosfati è che il trattamento con acqua causa la rottura delle catene e, di conseguenza, la conversione in prodotti di basso peso molecolare.

Le rocce, i mattoni, i calcestruzzi e le ceramiche sono per la maggior parte polimeri inorganici tridimensionali. I loro impieghi per molte applicazioni sono limitati per il fatto che essi sono difficili da trasformare in oggetti di largo impiego se non a elevate temperature. Essi non sono, in generale, flessibili, elastomerici o resistenti all'urto. Sebbene posseggano un'eccellente stabilità in aria alle alte temperature, essi non rispondono alle molteplici esigenze della moderna tecnologia.

Da un buon numero di anni i chimici si sono resi conto che una risposta a questo problema deve esistere nel campo intermedio tra i polimeri organici da una parte e i materiali inorganici di tipo minerale dall'altra. Sostanzialmente occorrono polimeri che abbiano una struttura lineare, gruppi sostituenti non ionici che favoriscano la flessibilità ed elementi inorganici nello scheletro che conferiscano stabilità al calore e all'ossidazione. La presenza di elementi inorganici nella catena polimerica conferirebbe inoltre flessibilità ed elasticità entro un ampio intervallo di temperatura. In teoria potrebbero presentarsi centinaia di strutture con differenti catene inorganiche. In pratica devono essere prese in considerazione alcune limitazioni. La prima restrizione è che gli elementi che costituiscono lo scheletro devono essere legati tra loro mediante legami covalenti. Questa limitazione viene considerata per ridurre la possibilità che le catene siano suscettibili di attacco da parte dell'acqua o che possano esistere legami incrociati ionici. Per questa ragione la ricerca principale per ottenere polimeri inorganici utilizzabili è concentrata sugli elementi covalenti principali in particolare boro, alluminio, silicio, germanio, stagno, azoto, fosforo, arsenico, antimonio, ossigeno, zolfo, selenio e tellurio. Uno dei primi polimeri inorganici a legami covalenti che siano stati ottenuti è lo zolfo polimerico. La forma stabile dello zolfo, a temperatura ambiente, è lo zolfo rombico, che contiene molecole cicliche con otto atomi di zolfo nell'anello (*si veda l'illustrazione in alto nella pagina a fronte*).

Lo zolfo rombico è una sostanza cristallina fragile, che fonde a 113 gradi centigradi per formare un liquido giallo-rosso. Lo zolfo liquido possiede la curiosa proprietà per cui al di sopra di 159 gradi centigradi la sua viscosità aumenta con l'aumentare della temperatura, contrariamente al comportamento di quasi tutti gli altri liquidi. L'aumento di viscosità deriva dall'apertura degli anelli a otto atomi con la conseguente conversione in lunghe molecole. Poiché le catene lunghe possono aggrovigliarsi piú vistosamente degli anelli, la viscosità aumenta. Quando la temperatura sale al di sopra di circa 175 gradi centigradi, tuttavia, la viscosità comincia a diminuire e questo indica che sta avvenendo una depolimerizzazione che ritrasforma le catene in anelli.

La forma altamente polimerica dello zolfo può essere « temprata » o isolata, versando la massa fusa a caldo in acqua. Quando è raffreddato, lo zolfo forma agglomerati gommosi o fibre (*si veda l'illustrazione a pagina 56*). Il polimero non è stabile a temperatura ambiente; indurisce lentamente non appena le lunghe catene subiscono il processo di cristallizzazione e si ritrasforma negli anelli ciclici a otto elementi. Per questa ragione lo zolfo polimerico non può trovare impiego pratico. Tuttavia il suo comportamento termico illustra un caratteristico comportamento di molti sistemi polimerici inorganici: gli anelli possono essere convertiti in catene polimeriche ad alto peso molecolare e a temperature elevate questo processo tende a invertirsi.

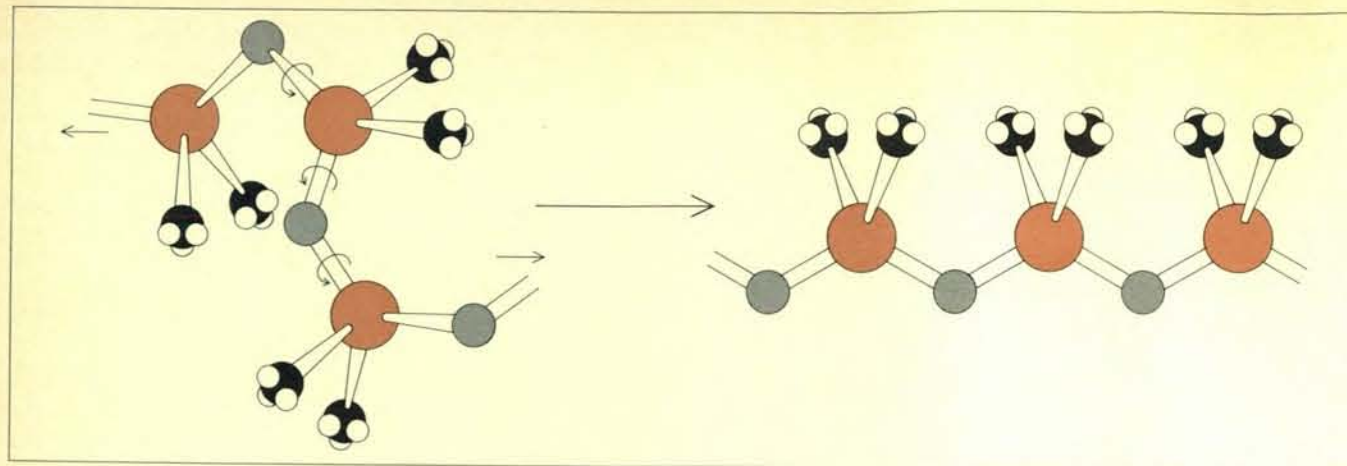
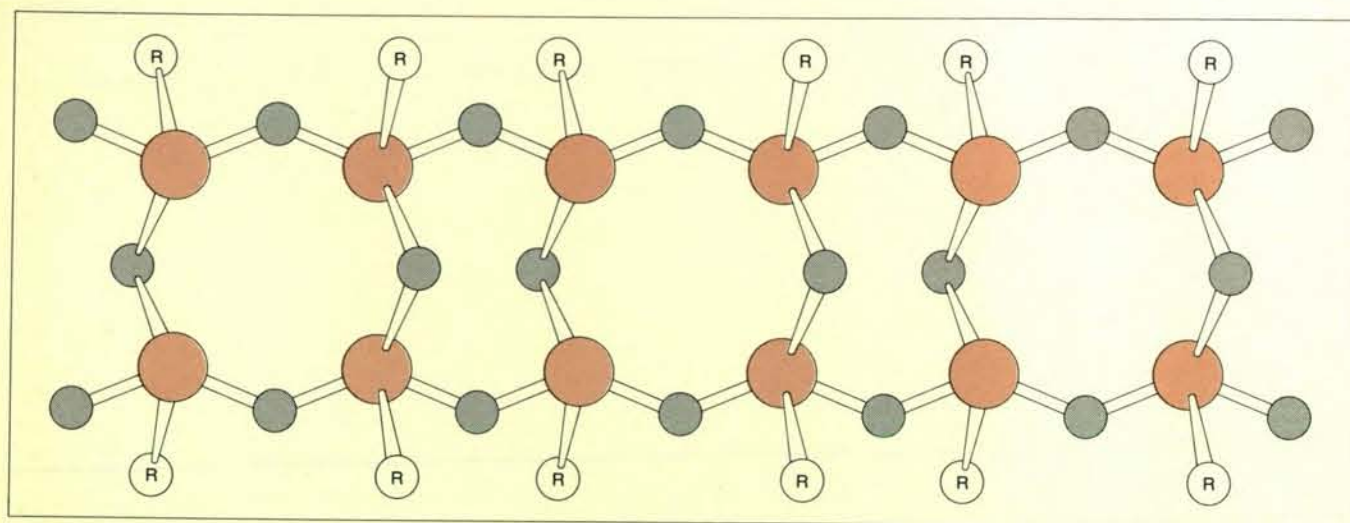
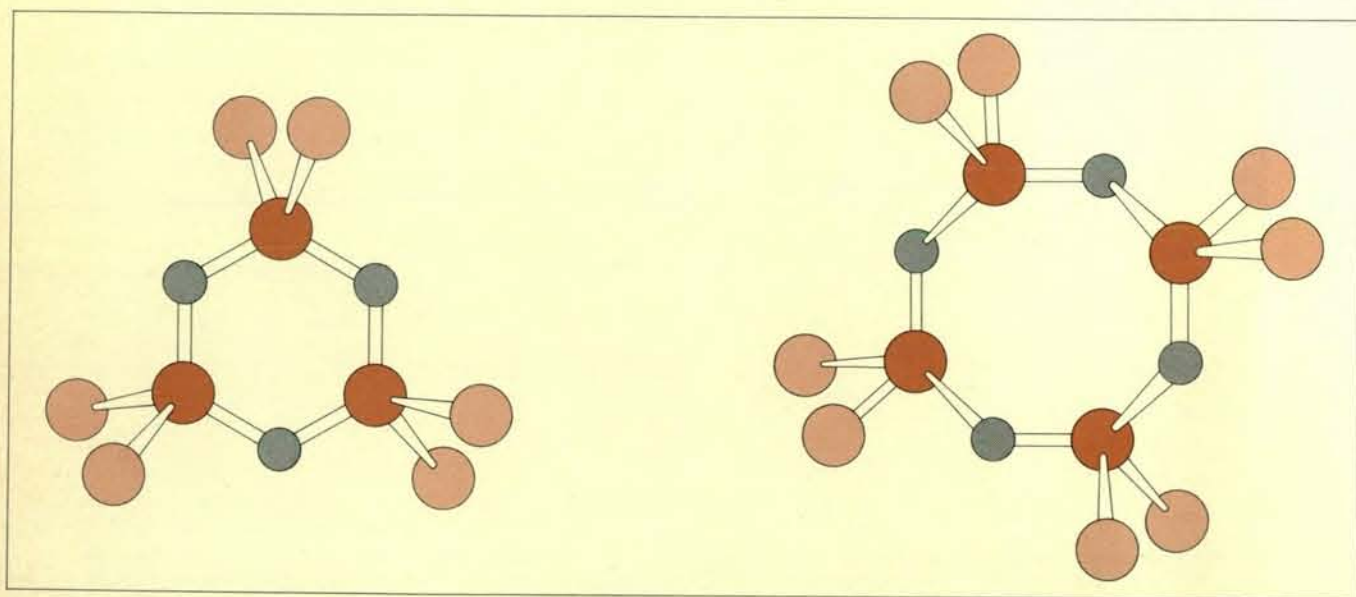
Quello che si vorrebbe sintetizzare è un polimero che rassomigli ai vetri inorganici, ma che sia modificato in modo che le catene polimeriche presentino flessibilità alle ordinarie temperature. Come abbiamo visto, la rigidità del vetro e dei polimeri similari è dovuta largamente alla presenza di gruppi laterali ionici. Se i gruppi laterali polari potessero essere sostituiti da unità prive di carica, la flessibilità delle catene dovrebbe risultare maggiore.

Questo concetto è l'idea che ha promosso la costruzione dei polimeri siliconici. Il silicone maggiormente impiegato, polidimetilsilossano, consiste in una catena in cui atomi di silicio si alternano ad atomi di ossigeno, con due gruppi metilici legati a ciascun atomo di silicio. Come lo zolfo polimerico, i siliconi sono ottenuti mediante polimerizzazione di un composto ciclico ad anello contenente otto atomi nell'anello stesso. Qui il materiale di partenza è una sostanza incolore, oleosa, nota come octametilciclotetrasilossano, la cui

POLIMERO AMORFO — POLIMERO CRISTALLINO
AUMENTO DI TEMPERATURA
LIQUIDO — LIQUIDO
TEMPERATURA DI FUSIONE
VISCOIDE — TERMOPLASTICO FLESSIBILE
GOMMA
TEMPERATURA DI TRANSIZIONE VETROSA — TEMPERATURA DI TRANSIZIONE VETROSA
VETRO — VETRO

Polimeri amorfi e cristallini mostrano proprietà differenti e distinte nell'intervallo di temperatura intermedio, quando sono troppo caldi per essere vetrosi e troppo freddi per essere liquidi. I polimeri amorfi, quando vengono riscaldati al di sopra della temperatura di transizione vetrosa, dapprima diventano gommosi, quindi viscosi ed infine liquidi, senza una netta distinzione tra le varie fasi. Al contrario, i polimeri cristallini sono flessibili e termoplastici in un intervallo di temperatura sufficientemente ampio.

I cristalli ordinari di zolfo rombico sono caratterizzati da anelli formati da otto unità di atomi di zolfo. Quando viene riscaldato al disopra di 140 °C, l'anello a otto atomi si apre e si lega ad altri anelli aperti per formare zolfo plastico polimerico (*destra*). Quando il polimero è raffreddato a temperatura ambiente, lentamente si ritrasforma nella forma ciclica. Il comportamento dello zolfo polimerico quando viene raffreddato rapidamente in acqua è illustrato dalla fotografia di pagina 56.

Il polimero siliconico (*destra*) viene ottenuto mediante polimerizzazione di octametilciclotetrasilossano (*a sinistra*) che è costituito da anelli formati da otto unità in cui quattro atomi di silicio (*in colore*) si alternano con quattro atomi di ossigeno. A ciascun atomo di silicio sono legati due gruppi metilici (CH_3). Il polimero lineare può essere trasformato in gomma siliconica mediante formazione di legami a ponte tra gruppi laterali. La reticolazione previene lo slittamento di una catena sull'altra quando il materiale viene teso. Il silicone è stato il primo polimero commerciale avente uno scheletro inorganico.

L'elasticità di un polimero quale la gomma siliconica è funzione della facilità con la quale le catene possono essere allungate a partire da una disposizione casuale a spirale. L'allungamento delle catene avviene per svolgimento dei legami stessi.

I polimeri siliconici con struttura a scala sono gli analoghi non ionici dei silicati anfiboli. Il polimero a doppia catena è più rigido di quello lineare. Qui *R* sta per un gruppo fenilico (C_6H_5), un anello benzenico privo di un atomo di idrogeno.

Due precursori di gomma inorganica sono gli anelli in cui atomi di fosforo (*in colore scuro*) si alternano con atomi di azoto (*grigio*): due atomi di cloro sono legati a ciascun atomo di fosforo. Gli anelli hanno formula generale $(NPCl_2)_n$ dove *n* è 3 o un numero intero più alto. Il membro più semplice della serie (*a sinistra*), ha un anello a sei atomi. Il tetramero (*a destra*) ha un anello a otto atomi. Per riscaldamento l'anello si apre e i composti polimerizzano formando il polidiclorofosfazene.

molecola ha un anello formato da quattro atomi di silicio alternati con quattro atomi di ossigeno. Quando l'anello a otto atomi viene riscaldato a una temperatura superiore ai 100 gradi centigradi, in presenza di tracce di un acido o di una base, esso polimerizza (*si veda l'illustrazione in basso nella pagina 49*). Il polimero finale può contenere almeno 30 000 unità silicio-ossigeno. I polimeri siliconici sono liquidi altamente viscosi o gommosi. Per unione delle catene polimeriche tra loro mediante un processo di reticolazione si ottengono gomme siliconiche. Questa reticolazione impedisce lo scorrimento di una catena sull'altra quando il materiale viene stirato.

Forse la caratteristica più sorprendente degli alti polimeri siliconici è che essi sono flessibili ed elastici entro un intervallo di temperatura straordinario. La loro temperatura di transizione vetrosa è –130 gradi centigradi. L'intervallo di temperatura di elasticità per le gomme siliconiche va da –60 gradi a 250 gradi; la temperatura più bassa indica l'inizio della cristallizzazione. La flessibilità della massa polimerica mette in risalto la facilità con cui i legami della struttura molecolare possono flettersi (*si veda l'illustrazione in alto nella pagina a fianco*). Per la verità i polimeri siliconici sono tra le più flessibili macromolecole conosciute.

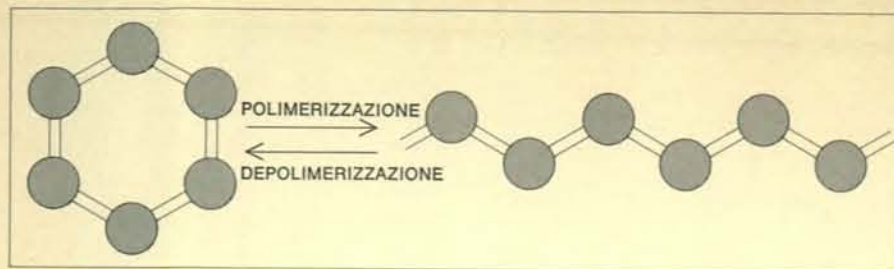
La loro elevata mobilità può essere attribuita alla mancanza di carica elettrica nei gruppi laterali e al fatto che tali gruppi sono legati ad atomi alterni dello scheletro molecolare e non a tutti gli atomi come avviene in molti polimeri organici. Quest'ultima caratteristica significa che vi sono ben poche opportunità per i gruppi laterali di venire in collisione tra loro o anche di attrarsi o respingersi quando i legami dello scheletro subiscono movimenti di flessione. È intuitivo che l'estrema flessibilità dello scheletro siliconico rende le gomme siliconiche altamente permeabili all'ossigeno. Per questo, film siliconici sono stati esperimentati in apparati per branchie artificiali che fossero in grado di estrarre l'ossigeno disciolto nell'acqua per impieghi sottomarini.

I siliconi sono anche fortemente idrorepellenti. In parte per questa loro proprietà, trovano impiego come prodotti lucidanti per automobili, in formulazioni antiadesive per batterie da cucina e in attrezzi biomedici. Per esempio, valvole cardiache artificiali e pompe sperimentali impiegate in cardiochirurgia, spesso sono fabbricate in gomma siliconica poiché il polimero ha una scarsa tendenza a favorire la coagulazione del sangue o a irritare i tessuti.

Come ho mostrato, una delle spinte allo sviluppo dei polimeri aventi scheletri inorganici, è stata la convinzione che queste sostanze sarebbero state più stabili a temperature elevate dei polimeri organici. Sebbene i siliconi siano certamente più resistenti all'ossidazione alle alte temperature, essi presentano uno svantaggio: a temperature superiori ai 250 gradi centigradi la catena silossanica si rompe per formare anelli, cosicché le vantaggiose proprietà del polimero vengono completamente perdute. La depolimerizzazione è simile a quella osservata con lo zolfo polimerico.

Un modo ovvio per risolvere questo problema è quello di costruire polimeri che assomiglino agli anfiboli, cioè silicati a catena doppia. Polimeri non ionici analoghi a questi sono stati preparati per la prima volta presso il Laboratorio di ricerca e sviluppo della General Electric, con gruppi fenilici al posto degli atomi di ossigeno polari. I materiali risultanti sono chiamati polimeri siliconici a « scala » (*si veda l'illustrazione al centro nella pagina a fianco*). Come ci si poteva aspettare, la struttura a doppia catena limita la mobilità dei legami silicio-ossigeno nello scheletro; i polimeri sono sostanze altofondenti, non elastomere. Quando vengono disciolti in solventi organici danno luogo a soluzioni viscose. Polimeri che rigonfiano per aggiunta di piccole quantità di solventi possono essere stirati e orientati. I polimeri fenilsiliconici reticolati « a scala » rimangono stabili a temperature superiori ai 300 gradi centigradi.

Nel nostro laboratorio presso l'Università di Stato della Pennsylvania abbiamo sintetizzato e studiato i componenti di un diverso gruppo di polimeri inorganici, probabilmente più versatile, denominati polifosfazeni. Questi polimeri hanno atomi di fosforo e di azoto che si alternano nel loro scheletro. È noto da almeno un secolo che due semplici composti inorganici, il pentacloruro di fosforo e il cloruro di ammonio reagiscono tra loro per formare una serie di composti ciclici inorganici. Tali prodotti hanno la formula generale $(NPCl_2)_n$ dove *n* è 3,4,5,6,7... I composti hanno una struttura ad anello formata da atomi di azoto e di fosforo alternati, con due atomi di cloro legati a ciascun atomo di fosforo. Il trimero (n = 3) e il tetramero (n = 4) sono i prodotti che si formano in maggior quantità (*si veda l'illustrazione in basso nella pagina a fianco*). Questi due composti sono solidi bianchi, stabili, che

Polifosfazeni sostituiti sono stati sintetizzati nel laboratorio dell'autore presso l'Università di Stato della Pennsylvania, nella ricerca di analoghi del polidiclorofosfazene resistenti all'azione dell'umidità. Il sistema consiste nel sostituire gli atomi di cloro con gruppi laterali organici più stabili. Questo può essere fatto trattando il polidiclorofosfazene (*in alto*) con sali sodici di alcooli o fenoli (*a, c*) nel qual caso il cloro viene eliminato come cloruro sodico, o con una ammina quale la metilammina (*b*), nel qual caso il cloro viene eliminato come acido cloridrico. Il prodotto della reazione *a* è un polimero termoplastico, cristallino, che forma un film; il prodotto della reazione *b* è un polimero solubile in acqua. Da sostituenti misti, secondo reazioni di tipo *c*, in cui vengono impiegati due differenti alcoolati di sodio (NaOR e NaOR′), derivano in genere elastomeri amorfi. Questi prodotti sono allo studio per scopi tecnologici.

La depolimerizzazione, cioè la conversione di una catena polimerica in un anello, pone un limite superiore alla temperatura alla quale molti polimeri derivati da composti inorganici ciclici possono essere impiegati. Sia la polimerizzazione che la depolimerizzazione riflettono la tendenza dei legami dei composti inorganici a essere distorti quando i composti sono sottoposti a elevate temperature. L'obiettivo del chimico dei polimeri è trovare i mezzi per aumentare la stabilità delle catene polimeriche.

fondono rispettivamente a 114 e 128 gradi centigradi e solubili in solventi organici. Quando i composti, allo stato fuso, vengono riscaldati a 250 gradi centigradi per parecchie ore, polimerizzano formando un alto polimero trasparente, gommoso. Il materiale risultante è conosciuto come «gomma inorganica» o polidiclorofosfazene. Poiché contiene nella catena almeno 15 000 unità fosforo-azoto, il suo peso molecolare è superiore a 2 000 000.

Considerando che il polidiclorofosfazene deriva esclusivamente da composti inorganici, esso può considerarsi un composto eccezionale. Nel suo comportamento tensione-rilassamento esso è un elastomero migliore della gomma naturale. Comunque, ancora piú interessante è il fatto che esso rimane gommoso quando viene raffreddato e indurisce solo quando la temperatura si abbassa a valori vicini alla temperatura di transizione vetrosa: –63 gradi centigradi. Cosí, come nel caso della gomma siliconica, catene di atomi inorganici possono dare polimeri con alta flessibilità della catena molecolare. Sfortunatamente il polidiclorofosfazene presenta una proprietà chimica che non consente il suo impiego come materia plastica comune. Esso reagisce lentamente con l'umidità atmosferica per dare acido fosforico, ammoniaca e acido cloridrico: l'elastomero cosí si sbriciola in una polvere. Era nostra idea che questa scissione idrolitica fosse iniziata da una reazione dei legami fosforo-cloro con acqua, piuttosto che da una diretta rottura della catena. Perciò abbiamo ipotizzato che se gli atomi di cloro fossero sostituiti da gruppi non idrolizzabili, la decomposizione del polimero sarebbe impedita. L'ostacolo piú grande è stata la necessità di escogitare un metodo per la sintesi di altri polimeri fosfazenici contenenti gruppi laterali organici.

Abbiamo tentato due approcci per la risoluzione del problema. Il primo era basato sul fatto che, se i trimeri ciclici clorofosfazenici possono essere polimerizzati per riscaldamento, ci si può aspettare che fosfazeni ciclici contenenti sostituenti organici si comportino in modo simile. Conseguentemente abbiamo preparato un numero di fosfazeni ciclici in cui gli atomi di cloro erano stati sostituiti da gruppi organici quali fenile, fenossi, fenilammino, etossi e trifluoroetossi. Nessuno di questi composti poté essere polimerizzato. Noi adesso siamo convinti che questo insuccesso sia dovuto a motivi sia meccanici sia termodinamici.

Il secondo approccio è stato piú fortunato. Esso prevedeva la ricerca di un metodo per la sostituzione diretta degli atomi di cloro nel polidiclorofosfazene con gruppi sostituenti organici. Una indagine approfondita ha mostrato che il polidiclorofosfazene ad alto grado di polimerizzazione può esistere in due forme: una struttura lineare o ramificata senza legami a ponte e una modificazione reticolata. La forma reticolata rigonfia in solventi organici quali benzene, ma non si scioglie. Pertanto questa forma non può essere usata per reazioni di sostituzione. La forma non reticolata si scioglie in solventi organici per formare soluzioni viscose. Mediante controllo accurato delle condizioni di polimerizzazione può essere isolato con alte rese il polimero non reticolato, ma la polimerizzazione prolungata porta a una pressoché totale conversione alla forma reticolata. La chiave del processo sta nell'eliminazione di questa reazione di reticolazione.

Soluzioni del polimero non reticolato possono essere trattate con reagenti in grado di sostituire il cloro. Quali reagenti possono essere impiegati l'etilato sodico, il trifluoroetilato sodico, il fenato sodico e varie ammine quali metilammina e anilina. Operando a condizioni di reazione appropriate tutti gli atomi di cloro possono essere sostituiti (*si veda l'illustrazione a pagina 51*).

I polimeri preparati con questo metodo sono risultati molto resistenti all'idrolisi; nel nostro laboratorio abbiamo sintetizzato piú di 40 polimeri differenti. In funzione del tipo di gruppo laterale introdotto, i polimeri variano da elastomeri a bassa temperatura a termoplastici flessibili, formanti film, o vetri trasparenti. Gruppi sostituenti diversi conferiscono ai polimeri proprietà superficiali differenti, solubilità selettive e, in certi casi, una resistenza ai solventi inaspettata. Infatti oggi sembra possibile conferire al polimero tutta una serie di requisiti mediante una scelta giudiziosa dei gruppi sostituenti. Questa versatilità dei sistemi polimerici fosfazenici costituisce la loro caratteristica piú notevole.

Due esempi possono servire a illustrare questo punto. Quando il polidiclorofosfazene reagisce con il trifluoroetilato sodico, gli atomi di cloro sono sostituiti da gruppi trifluoroetossi– (CF_3CH_2O–). Il polimero è chiamato polibis(trifluoroetossi)fosfazene, e ha la formula $[NP(OCH_2CF_3)_2]_n$. Può essere ottenuto sotto forma di film incolore, opalescente, flessibile che, in apparenza, assomiglia superficialmente al film di polietilene. Soluzioni di polimero possono essere estruse in un mezzo non solvente per dare fibre flessibili e leggermente elastiche. Il polimero ha una bassa temperatura di transizione vetrosa (–66 gradi centigradi) e rimane flessibile da questa temperatura fino al suo punto di fusione a 242 gradi centigradi. La presenza dei gruppi fluorurati laterali rende il polimero altamente idrorepellente, piú del Teflon o dei siliconi.

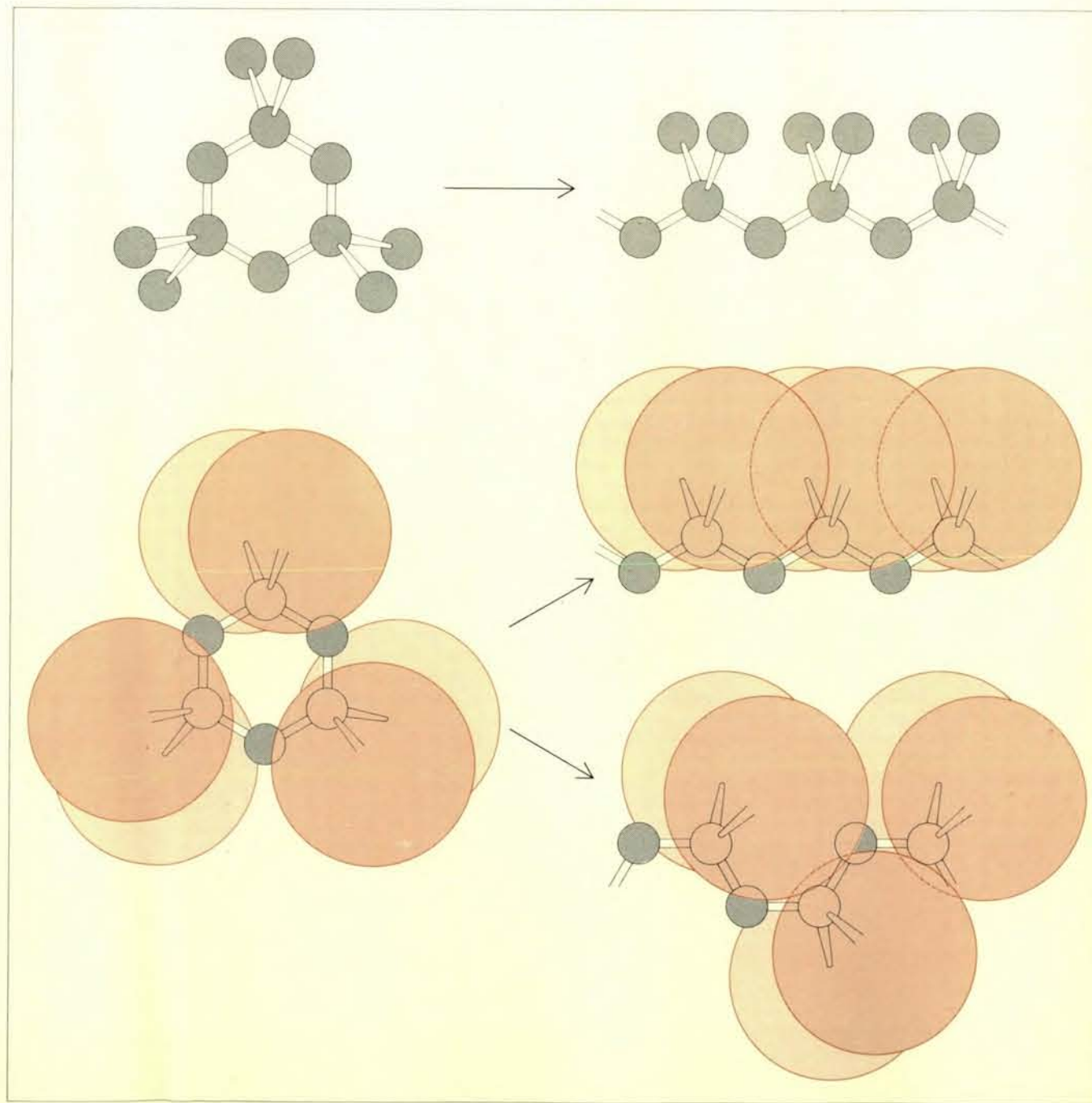
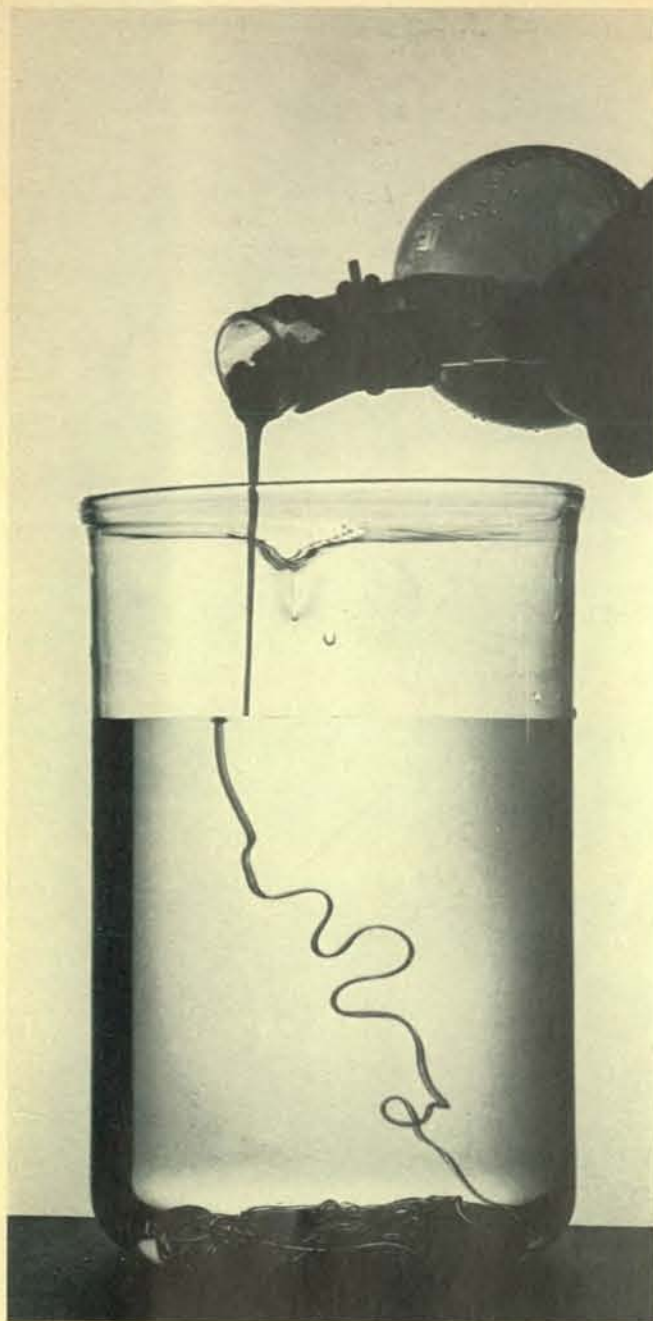
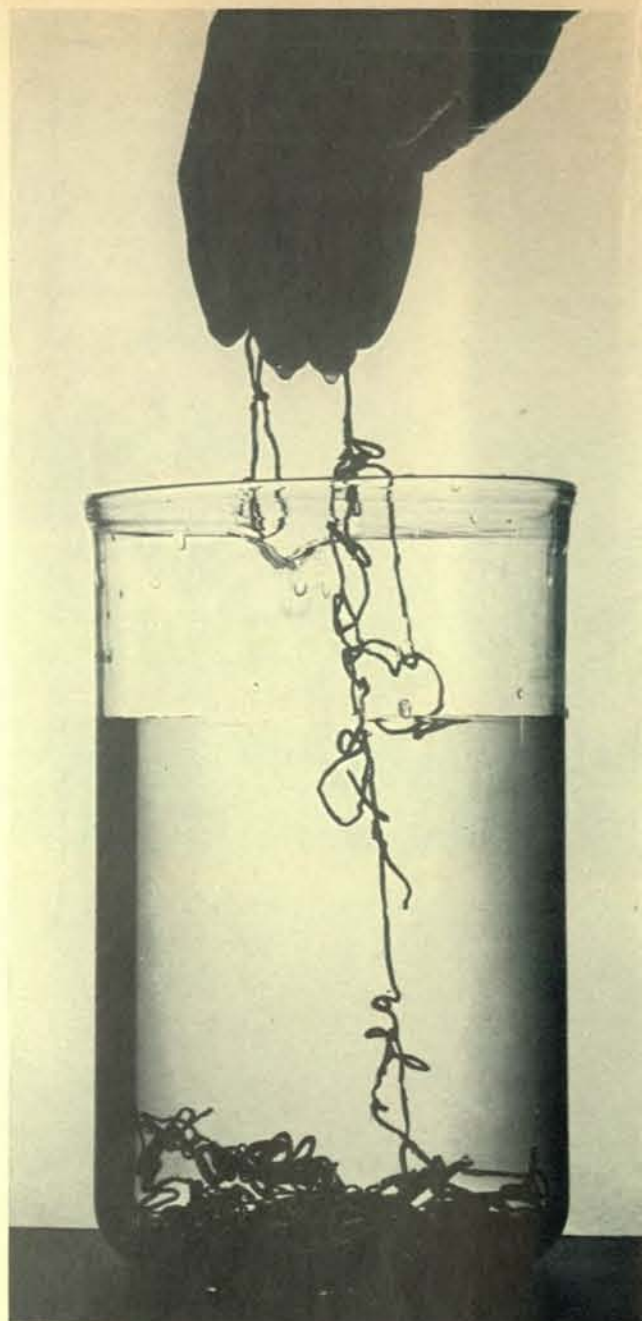
L'aspetto opalescente dei film di questo polimero e anche le sue proprietà termoplastiche sono una conseguenza della sua alta cristallinità, dovuta alla regolare disposizione dei segmenti di catene adiacenti. I polimeri che non sono cristallini tendono a essere elastomeri al di sopra della loro temperatura di transizione vetrosa. Se invece sono cristallini, formano film o fibre tenaci e flessibili.

La cristallinità di un polimero può generalmente essere distrutta dall'introduzione casuale di due o piú gruppi sostituenti diversi nella sua struttura. Se entrambi i gruppi trifluoroetossi e, ad esempio, eptafluorobutossi, sono legati alla catena fosfazenica, il polimero si comporterà come un elastomero gommoso, piuttosto che come una sostanza termoplastica flessibile.

Elastomeri di questo tipo sono tuttora allo studio presso la società Firestone Tire & Rubber per l'impiego come guarnizioni, anelli di tenuta, condotte di carburante a bassa temperatura, e in altre applicazioni per le quali sia richiesta resistenza ai solventi.

Un secondo esempio di struttura polifosfazenica è fornito dal polimero preparato per reazione di polidiclorofosfazene con metilammina. Il prodotto è una sostanza chiara, trasparente, termoplastica capace di formare film, avente la formula $[NP(NHCH_3)_2]_n$. Questo composto differisce da quasi tutti gli altri polimeri sintetici essendo solubile in acqua. I gruppi sostituenti amminici tendono a conferire proprietà idrofile piuttosto che idrofobe.

I poliorganofosfazeni, come i siliconi, presentano un inconveniente. Sopra i 200 gradi centigradi le catene del polimero si rompono per dare anelli piccoli e particolarmente per formare trimeri e tetrameri ciclici. Poiché le proprietà di impiego di un polimero dipendono quasi interamente dalla presenza di catene lunghe, un tale processo di depolimerizzazione è altamente svantaggioso. La depolimerizzazione pone un serio limite superiore alle temperature alle quali molti polimeri inorganici possono essere impiegati. L'arresto della depolimerizzazione termica costituisce uno dei principali obiettivi della ricerca nel campo dei polimeri inorganici.

La rottura delle catene dei polimeri inorganici con la formazione di anelli è la conseguenza della tendenza di molti composti inorganici a subire una «distorsione» dei loro legami a temperature elevate. Possiamo anche dire che la polimerizzazione degli anelli

La stabilità della catena è favorita dalla presenza di gruppi laterali di piccole dimensioni (*in alto*) che non interferiscono l'un l'altro quando un anello è polimerizzato in una catena. Se, al contrario, un anello contiene gruppi laterali voluminosi (*in basso*), la polimerizzazione diventa difficile poiché i gruppi laterali interferiscono fortemente. Inoltre, catene con gruppi laterali voluminosi tendono a depolimerizzare a temperature piú basse rispetto a catene con gruppi di piccole dimensioni.

Zolfo polimerico può essere preparato riscaldando il normale zolfo rombico a una temperatura tra 140 e 170 gradi C. Se il polimero è raffreddato rapidamente versandolo in acqua fredda (*a sinistra*), **esso manterrà le proprietà elastiche e la flessibilità per breve tempo a temperatura ambiente** (*a destra*). **Alla fine il polimero ritornerà alla sua struttura originale ad anello.**

e la depolimerizzazione delle catene sono semplicemente aspetti differenti dello stesso processo di distorsione dello scheletro molecolare (*si veda l'illustrazione a pagina 54*). La domanda che ci si deve porre è questa: quale struttura molecolare in differenti sistemi inorganici può rendere stabili gli anelli piuttosto che le catene o viceversa? Se conoscessimo la risposta saremmo in grado di comprendere perché alcuni anelli inorganici non possono essere polimerizzati e perché alcune catene inorganiche sono termicamente piú stabili di altre.

Una gran quantità di fattori sono legati attualmente a questa domanda, ma tre considerazioni sono particolarmente importanti. Consideriamo dapprima alcuni composti ad anello organici ed inorganici che non hanno potuto essere polimerizzati in catene lineari. Per esempio il benzene (C_6H_6) non può essere polimerizzato; infatti l'anello benzenico è una delle unità piú stabili conosciute nella chimica. Analogamente le *s*-triazine, le borazine e i composti ciclici zolfo-azoto conosciuti come alogenuri tiazilici e sulfanurici, resistono alla polimerizzazione. Una caratteristica che tutti questi sistemi ciclici hanno in comune è che essi sono stabilizzati a un livello piú o meno elevato da una configurazione elettronica « aromatica ». Questa configurazione è caratteristica degli elettroni nel benzene e nei sistemi ciclici a esso correlabili. Quindi non vi è tendenza a formare catene per apertura degli anelli. Gli alti polimeri lineari apparentemente non possono essere stabilizzati per questo effetto.

Un secondo fattore legato al problema è connesso con le dimensioni dei gruppi laterali legati alla catena principale. Come abbiamo appena visto, i clorofosfazeni ciclici polimerizzano formando catene lunghe, mentre gli anelli con sostituenti organici non hanno potuto essere finora polimerizzati. Inoltre, gli esperimenti hanno mostrato che quando i polimeri organofosfazenici sono preparati con l'altro sistema sopra descritto, essi depolimerizzano, formando anelli, a una temperatura piú bassa rispetto agli alti polimeri clorofosfazenici.

La tendenza per certi polimeri a depolimerizzare dipende in larga misura dalle dimensioni dei gruppi laterali, poiché questi gruppi possono interferire piú liberamente tra loro quando il composto è in forma di catena piuttosto che quando è in forma di anello (*si veda l'illustrazione a pagina 55*).

Se i gruppi laterali sono unità organiche di grandi dimensioni, entreranno in collisione o si respingeranno l'un l'altro fortemente nel polimero, mentre si eviteranno completamente in una molecola ciclica piccola. Cosí, la conversione di un anello ciclico trimero in un polimero sarà difficile; per converso, se il polimero può essere preparato per qualche altra via indiretta, sarà meno stabile degli anelli. Data una opportunità, esso si scinderà per formare anelli dato che questo processo facilmente porta il sistema a un piú basso livello energetico. Questa è probabilmente una delle piú importanti ragioni per cui i polimeri organofosfazenici e alcuni polimeri organosilossanici depolimerizzano alle alte temperature.

Un terzo fattore importante nella depolimerizzazione di polimeri inorganici e anche di molti polimeri organici è legata all'entropia del sistema. L'entropia è il grado di disordine del sistema e (particolarmente alle alte temperature) i sistemi cambiano per generare il piú alto grado di disordine. La rottura di una molecola di polimero in, per esempio, 10 000 anelli è un processo che aumenta notevolmente il disordine del sistema. Conseguentemente le alte temperature generalmente favoriscono la conversione di catene lunghe in piccoli anelli.

Questi fatti dimostrano che la sintesi di polimeri inorganici che presentino un'utilità pratica è un problema complesso ed eccitante. Inoltre quegli alti polimeri che sono stati sintetizzati sono spesso cosí sorprendentemente differenti dalla maggior parte dei polimeri organici, che la ricerca per la sintesi di nuove macromolecole inorganiche continuerà senz'altro. Al momento attuale questo campo si trova ancora in uno stadio infantile, con molti elementi e letteralmente centinaia di combinazioni elementari che devono ancora essere esplorate.

Tettonica a zolle e vita negli oceani

La frattura dell'antico supercontinente di Pangea e l'allontanamento dei suoi frammenti hanno innescato un processo evolutivo a lungo termine che ha condotto alla varietà senza precedenti dell'attuale biosfera

di James W. Valentine e Eldridge M. Moores

Una profonda rivoluzione concettuale ha sconvolto le scienze della Terra negli anni Sessanta. La nuova visione del mondo ha modificato radicalmente nozioni acquisite ormai da lunga data, come quella della permanenza dei continenti e dei bacini oceanici, e ha permesso di afferrare le cause profonde e il significato della maggior parte delle principali strutture del mantello e della crosta terrestri. Come conseguenza di questo processo conoscitivo è ormai accettato quasi da tutti il fatto che, nel corso della storia geologica, la posizione geografica, la distribuzione e addirittura le dimensioni e il numero dei continenti siano mutati notevolmente. Questi processi di deriva, di frammentazione e riunione delle masse continentali si sono sviluppati per almeno 700 milioni di anni e, forse, per oltre 2 miliardi di anni.

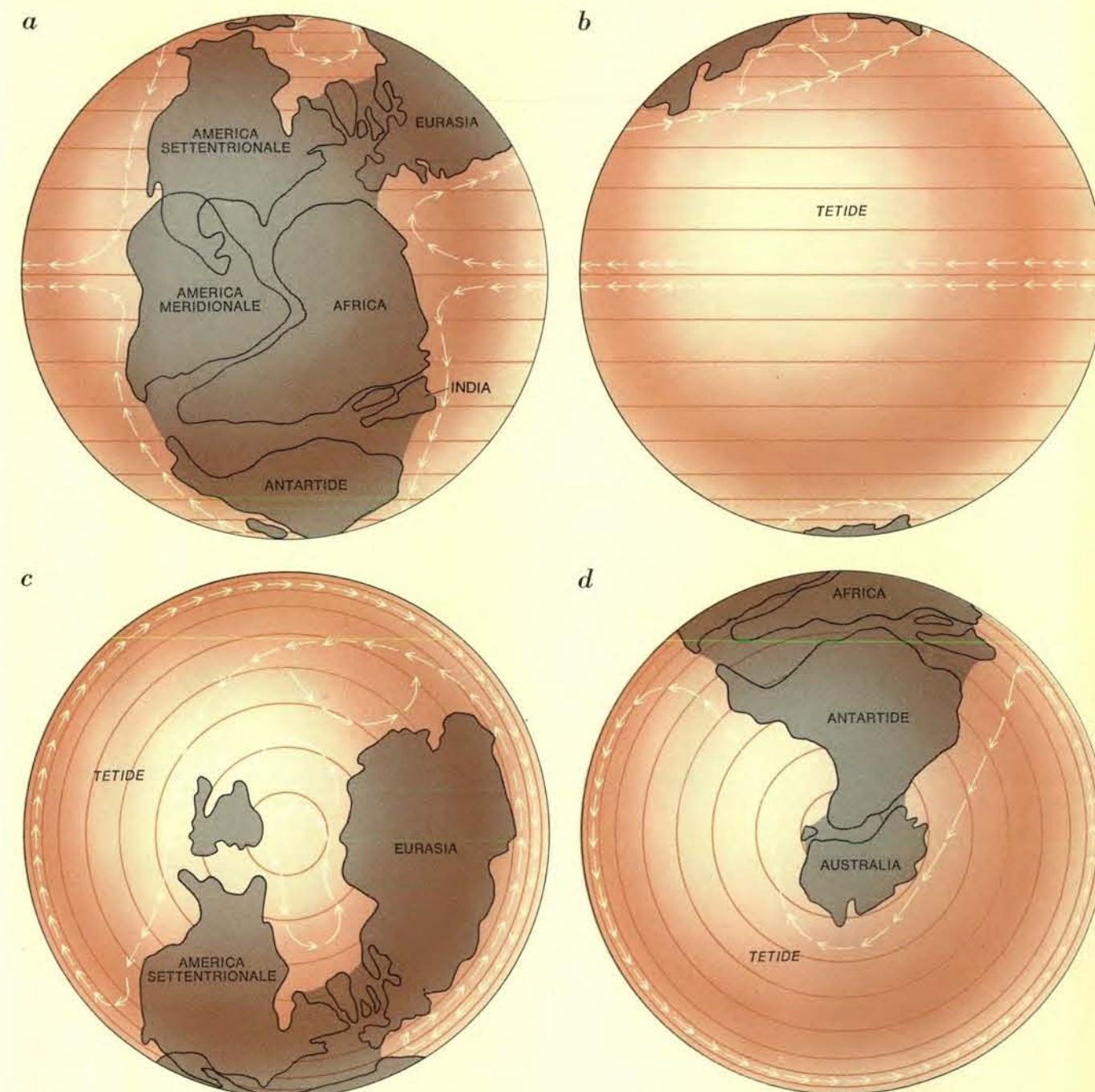
Naturalmente modificazioni di così grande portata sulla configurazione relativa dei continenti e degli oceani devono avere provocato profondi effetti sull'ambiente mutando la distribuzione dei climi e influenzando di conseguenza la composizione e la distribuzione della vita nella biosfera. Tali cambiamenti più o meno continui nell'ambiente devono avere influito sensibilmente anche sul corso stesso dell'evoluzione e perciò anche sulla storia stessa della vita.

La selezione naturale, il meccanismo fondamentale attraverso cui procede l'evoluzione, è un processo estremamente complicato. Pur con i limiti imposti dal meccanismo dell'ereditarietà, la selezione naturale è sostanzialmente un fenomeno ecologico basato sui rapporti tra gli organismi e l'ambiente in cui vivono. Per ciascuna specie, determinati caratteri ereditabili sono favoriti proprio perché più adatti alla sopravvivenza e alla riproduzione in quel certo ambiente nel quale si manifestano. Per comprendere la ragione per cui un certo gruppo di organismi si è evoluto è necessario dunque conoscere due fattori principali. Primo, è necessario sapere che caratteristiche avesse l'organismo ancestrale che costituisce il « materiale grezzo » sul quale ha agito la selezione. Secondo, è necessario avere qualche idea della sequenza di condizioni ambientali che hanno orientato l'evoluzione del ceppo originario verso le forme successive. Dati questi due fattori, è possibile dedurne le interazioni tra organismo e ambiente che hanno determinato lo evento evolutivo. Lo studio delle relazioni tra gli antichi organismi e il loro ambiente è chiamato paleoecologia.

La nuova concezione di deriva dei continenti, emersa negli anni Sessanta, è strettamente legata alla teoria della tettonica a zolle. Secondo questa teoria, in alcuni centri di espansione posti al di sotto delle dorsali medio-oceaniche e in piccoli bacini oceanici, nuova crosta oceanica proveniente dal sottostante mantello va continuamente ad aggiungersi alla crosta terrestre, a una velocità che può raggiungere e superare i 10 cm all'anno. Il fondo oceanico si espande lateralmente allontanandosi da tali centri e, da ultimo, sprofonda nell'interno della Terra in corrispondenza di zone di subduzione che sono segnalate dalla presenza di profonde fosse oceaniche; i processi di assorbimento creano vulcani che si allineano in prossimità delle fosse. La litosfera, cioè l'involucro roccioso più esterno della Terra, dunque, comprende diverse zolle principali che si generano in corrispondenza dei centri di espansione e si consumano in corrispondenza delle zone di subduzione. La maggior parte delle zolle litosferiche reca sul dorso uno o più continenti che si muovono passivamente con la zolla su cui giacciono. Dato che i continenti sono troppo leggeri per sprofondare nelle fosse, essi rimangono in superficie. I continenti possono tuttavia spezzarsi in corrispondenza di nuove dorsali e ne deriva che tra un frammento e l'altro possono comparire bacini oceanici. All'inverso, i continenti possono essere condotti l'uno verso l'altro e saldarsi tra loro quando entrano in collisione presso una fossa oceanica. In decine o centinaia di milioni di anni interi bacini oceanici possono nascere o scomparire e il numero, le dimensioni e la distribuzione dei continenti possono mutare radicalmente.

Le tracce di tali fenomeni di frammentazione e riunione di continenti sono costituite dalle regioni deformate che caratterizzano le fasce orogeniche e soprattutto dalle fasce montuose che contengono quelle particolari formazioni rocciose note col nome di ofioliti. Tali formazioni sono caratterizzate da una precisa sequenza di rocce costituita (dalla base al tetto) di rocce ultramafitiche (rocce ricche di magnesio e costituite soprattutto di olivina), da gabbri (una roccia basaltica a grana grossolana), da rocce vulcaniche e rocce sedimentarie. Si ritiene che le principali fasce ofiolitiche della Terra rappresentino frammenti fossili di bacini oceanici ormai scomparsi (*si veda la figura alle pagine 60-61*). L'esistenza di tali fasce all'interno di un continente (per esempio, la fascia ofiolitica degli Urali in URSS), costituisce un indizio della presenza di un antico bacino oceanico che separava due frammenti continentali i quali, in un certo momento del passato, sono entrati in collisione tra loro e si sono saldati a costituire un continente unico di maggiori dimensioni. La geologia di un certo sistema montuoso permette di « leggere » la scala cronologica secondo la quale si sono verificati eventi come l'apertura dei bacini oceanici, la dispersione dei continenti e la chiusura degli oceani per collisione continentale.

Naturalmente, anche l'ambiente biologico subisce continue modifiche. Per esempio, i mutamenti nella distribuzione dei continenti influenzano profondamente le correnti oceaniche, la temperatura, il tipo di fluttuazioni stagionali, la distribuzione degli elementi nutrienti, l'andamento della produttività e molti altri fattori di fondamentale importanza per gli organismi viventi. Perciò, l'orientamento dei processi evolutivi degli organismi animali marini deve avere subito variazioni nel tempo geologico in seguito ai più importanti mutamenti ambientali e nella misura in cui la selezione naturale ha agito per favorire l'adattamento degli organismi alle nuove condizioni.

Il modello della tettonica a zolle fornisce oggi un quadro concettuale coerente nel quale inserire la successione dei mutamenti faunistici testimoniati dalla paleontologia. I paleontologi hanno infatti da tempo compreso che la documentazione paleontologica costituisce una buona testimonianza di profondi mutamenti nella composizione, distribuzione e diversificazione della vita nei mari. Oggi, tuttavia, per la prima volta è possibile ricostruire la sequenza dei mutamenti ambientali nel

Andamento probabile delle correnti oceaniche attorno al continente di Pangea, il supercontinente unico che si ritiene sia esistito all'inizio del Triassico, circa 225 milioni di anni fa; nelle figure che rappresentano due proiezioni equatoriali (*a*, *b*) e due proiezioni polari (*c*, *d*) le correnti sono indicate dalle frecce. Tenendo conto della combinazione di fattori geografici e ambientali, ivi comprese le correnti prevalentemente calde qui indicate, sarebbe lecito aspettarsi che il margine continuo ricoperto di acque basse che attorniava il supercontinente, fosse popolato da poche specie estremamente diffuse. I fossili di quel periodo indicano proprio un basso grado di diversificazione delle specie oltre a un grado di provincialità assai ridotto.

quadro della tettonica a zolle, analizzare le conseguenze paleoecologiche di questo modello e tentare di correlarle con la serie di mutamenti faunistici testimoniati dalla paleontologia. Tale completa ricostruzione può alla fine condurre alla spiegazione della maggior parte di quei cambiamenti faunistici che, pur essendo noti da molto tempo, sono rimasti, per ciò che attiene alle cause, un mistero. Anche in questo primo stadio del loro lavoro di reinterpretazione, i paleontologi sono riusciti a chiarire un certo numero delle piú importanti estinzioni e diversificazioni manifestatesi nel passato.

Il primo obiettivo da raggiungere per chiarire le relazioni che esistono tra tettonica a zolle e storia della vita è quello di comprendere i rapporti che esistono oggi tra la vita nei mari, l'attuale distribuzione dei continenti e il modello della tettonica a zolle. La grande maggioranza delle specie marine (circa il 90%) vive sulle piattaforme continentali, o in acqua bassa sulle porzioni sommerse di isole o su rilievi sottomarini la cui sommità si trova a profondità inferiori ai 200 metri: anche la maggior parte dei resti fossili è costituita da faune di questo tipo. Ne consegue che, ai nostri fini, è di particolare interesse stabilire i caratteri delle faune marine bentoniche (che vivono cioè sul fondo) di piccola profondità.

Le acque poco profonde piú ricche di animali si trovano oggi alle basse latitudini tropicali dove le comunità animali sono associate a un gran numero di specie altamente specializzate. Spostandosi verso latitudini piú elevate, diminuisce progressivamente il grado di diversificazione; nell'Artico o nell'Antartico vive, quando si considerino regioni comparabili, meno di un decimo degli animali che si sviluppano ai tropici (*si veda la figura alle pagine 62-63*). Il gradiente di diversificazione si correla bene con quello della stabilità della disponibilità di cibo; le fluttuazioni nella produttività primaria diventano piú elevate quando piú pronunciata si fa la stagionalità. Anche se l'andamento d'insieme della diversificazione delle faune terrestri è dominato da questo sensibile gradiente latitudinale, vi sono anche importanti tendenze alla diversificazione nel senso della longitudine. Per esempio, in regioni che si trovano a latitudini analoghe, la diversificazione è meno pronunciata laddove si registrano forti sbalzi stagionali (come le variazioni nell'andamento delle correnti di superficie o nella risalita di acque fredde) che modificano l'afflusso di elementi nutrienti e determinano di conseguenza notevoli fluttuazioni nella produttività.

- PRECAMBRIANO
- FASCIA PANAFRICANO-BAIKALIANA
- FASCIA CALEDONIANA
- FASCIA APPALACHIANO-ERCINICA
- FASCIA URALIANA
- FASCIA CORDILLERANO-TETICA

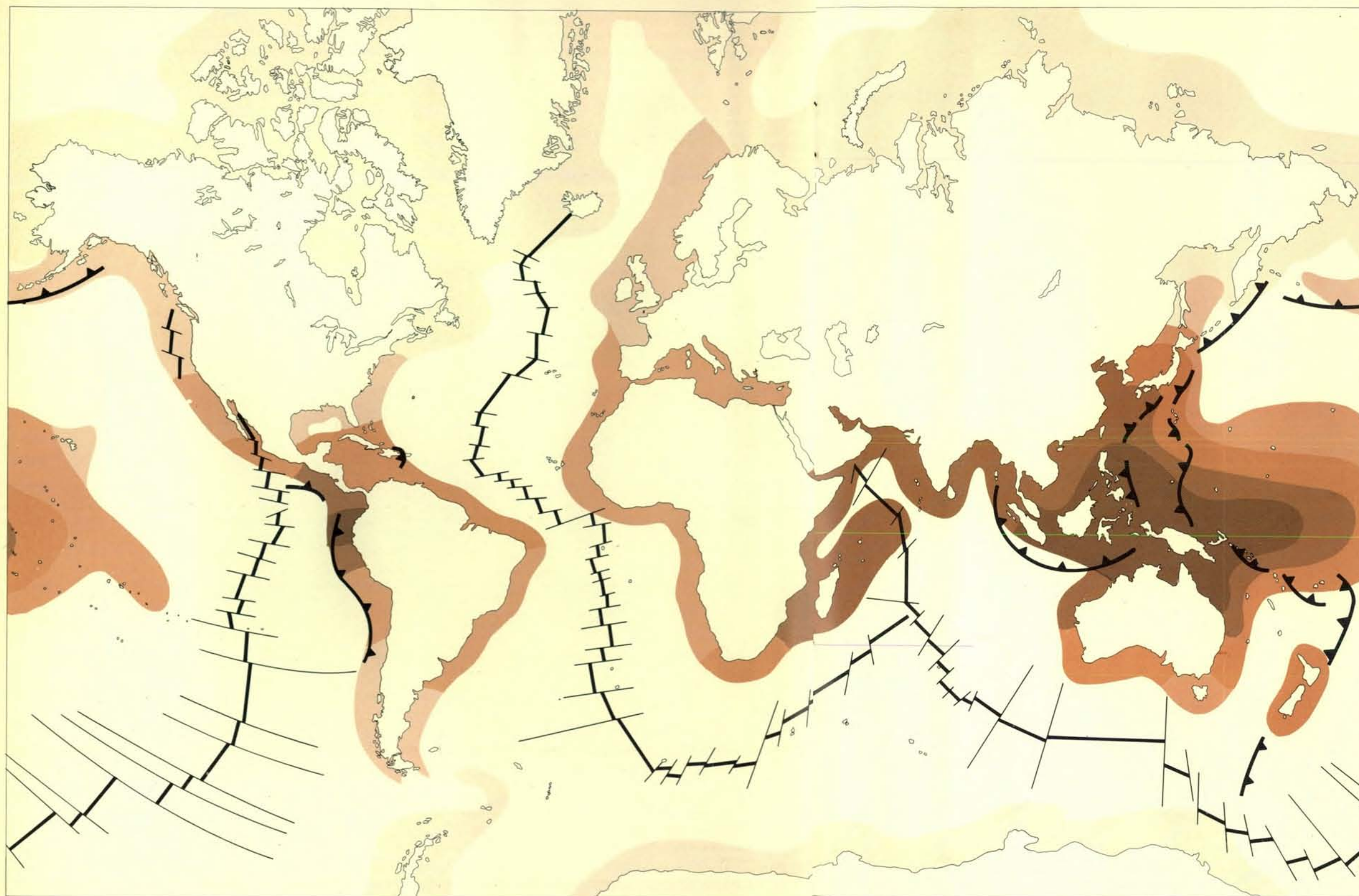
Le tracce geologiche dell'attività delle zolle nel passato sono conservate in alcune fasce montuose (*in colore*), e soprattutto in quelle che contengono le caratteristiche sequenze di rocce note col nome di ofioliti (*punti neri*). La fascia panafricano-baikaliana, per esempio, è costituita da rocce di età compresa tra 873 e 450 milioni di anni fa e può rappresentare la sutura di tutte o quasi tutte le terre emerse attorno agli inizi del Fanerozoico. Tale supercontinente può allora essersi spezzato in quattro o piú continenti piú piccoli poco prima dell'inizio e nel corso stesso del Cambriano. Il sistema caledoniano può rappresentare la collisione di due continenti avvenuta verso la fine del Siluriano o agli inizi del Devoniano (circa 400 milioni di anni fa). Il sistema appalachiano-ercinico può rappresentare una collisione fra due continenti avvenuta nel corso del Carbonifero (300 milioni di anni fa). I monti Urali possono rappresentare una collisione simile avvenuta all'incirca nel Permotriassico (220 milioni di anni fa). Il sistema cordillerano-tetico rappresenta regioni orogeniche mesozoiche e comprende le collisioni continentali che hanno avuto come risultato il sistema montuoso alpino-himalayano. Le fasce ofiolitiche qui illustrate sono ciò che resta di fondi oceanici portati ad affiorare dall'edificazione di tali sistemi montuosi. Le dorsali oceaniche sede di attiva espansione, come la Dorsale Medio-atlantica, sono indicate da linee nere spesse, interrotte da linee piú sottili che corrispondono a faglie trasformi. Le zone di subduzione sono rappresentate dalla serie di linee curve spesse e dentate.

A qualunque latitudine, tuttavia, la diversificazione è piú elevata al largo delle spiagge di piccole isole o di piccoli continenti che si trovano in vasti oceani; qui infatti le fluttuazioni nell'afflusso di elementi nutrienti sono meno influenzate dalle variazioni stagionali determinate dalla presenza di vaste aree emerse. La diversificazione è invece meno pronunciata attorno a vasti continenti, soprattutto quando questi si affacciano su piccoli oceani: le variazioni stagionali nelle acque basse sono piú intense. In breve, mentre la diversificazione latitudinale aumenta in genere dalle alte latitudini verso quelle basse, la diversificazione longitudinale aumenta in genere con la distanza da vaste aree di terre emerse. In entrambe queste tendenze l'aumento nella diversificazione è legato a un corrispondente incremento nella stabilità dell'afflusso di risorse alimentari. La distribuzione geografica di fonti stabili di risorse alimentari dipende largamente dalla forma dei continenti e potrebbe essere anche sensibile all'estensione di mari interni e alla presenza di catene montuose costiere. I mari che giacciono su piattaforme continentali rivestono notevole importanza: non soltanto le loro acque basse costituiscono un habitat ideale per le faune di scarsa profondità ma tali mari tendono anche a ridurre le differenze stagionali nel clima e ad avere un'influenza benefica sull'ambiente locale.

Le faune marine di scarsa profondità tendono oggi ad avere un elevato grado di provincialità; ciò significa che specie che vivono in oceani diversi o sulle opposte sponde dello stesso oceano tendono ad essere abbastanza diverse. Persino lungo linee di costa continue si registrano importanti mutamenti nella composizione delle associazioni faunistiche da luogo a luogo, variazioni queste che corrispondono in genere a mutamenti climatici. Il fondo profondo degli oceani, che si genera in corrispondenza delle dorsali oceaniche, costituisce una barriera importante nei confronti della diffusione degli organismi di acque basse; altre barriere sono costituite evidentemente dalla variazione climatica nel senso della latitudine. L'attuale serie di dorsali disposte soprattutto in senso nord-sud costituisce una struttura per la quale continenti e oceani si alternano nel sen-

so della longitudine a costituire una serie di barriere alla diffusione degli organismi marini di scarsa profondità. D'altro canto il netto gradiente climatico latitudinale crea una catena di province lungo le linee di costa orientate in senso nord-sud. Ne risulta che le faune marine attuali sono oggi suddivise in piú di 30 province faunistiche, fra le quali vi è in genere una ben piccola percentuale di specie comuni (*si veda la figura alle pagine 64-65*). Si stima che la fauna marina d'acqua bassa sia rappresentata da specie oltre 10 volte piú numerose di quelle che si osserverebbero in un mondo costituito da una sola provincia biogeografica anche altamente diversificata.

Gli archi vulcanici che compaiono al di sopra delle zone di subduzione formano catene pressoché continue di isole e costituiscono un'eccellente via di diffusione. Quando lunghe catene insulari sono disposte in senso est-ovest in modo tale da trovarsi all'interno della stessa zona climatica, esse sono abitate da faune notevolmente comprensive che sono assai differenziate per la loro latitudine. In effetti, la provincia marina piú comprensiva e contemporaneamente piú diversificata è quella indopacifica, le cui regioni centrali sono costituite da archi insulari. La vita delle faune di questa provincia si diffonde da questi archi verso le piattaforme continentali a occidente (India e Africa orientale) e verso le zone vulcaniche all'interno delle zolle (le isole della Polinesia e della Micronesia) che sono abbastanza vicine ad essi. Questa grande associazione faunistica tropicale è separata dalla America occidentale dalla Barriera Est-pacifica, un diaframma zoogeografico costituito da una dorsale in espansione.

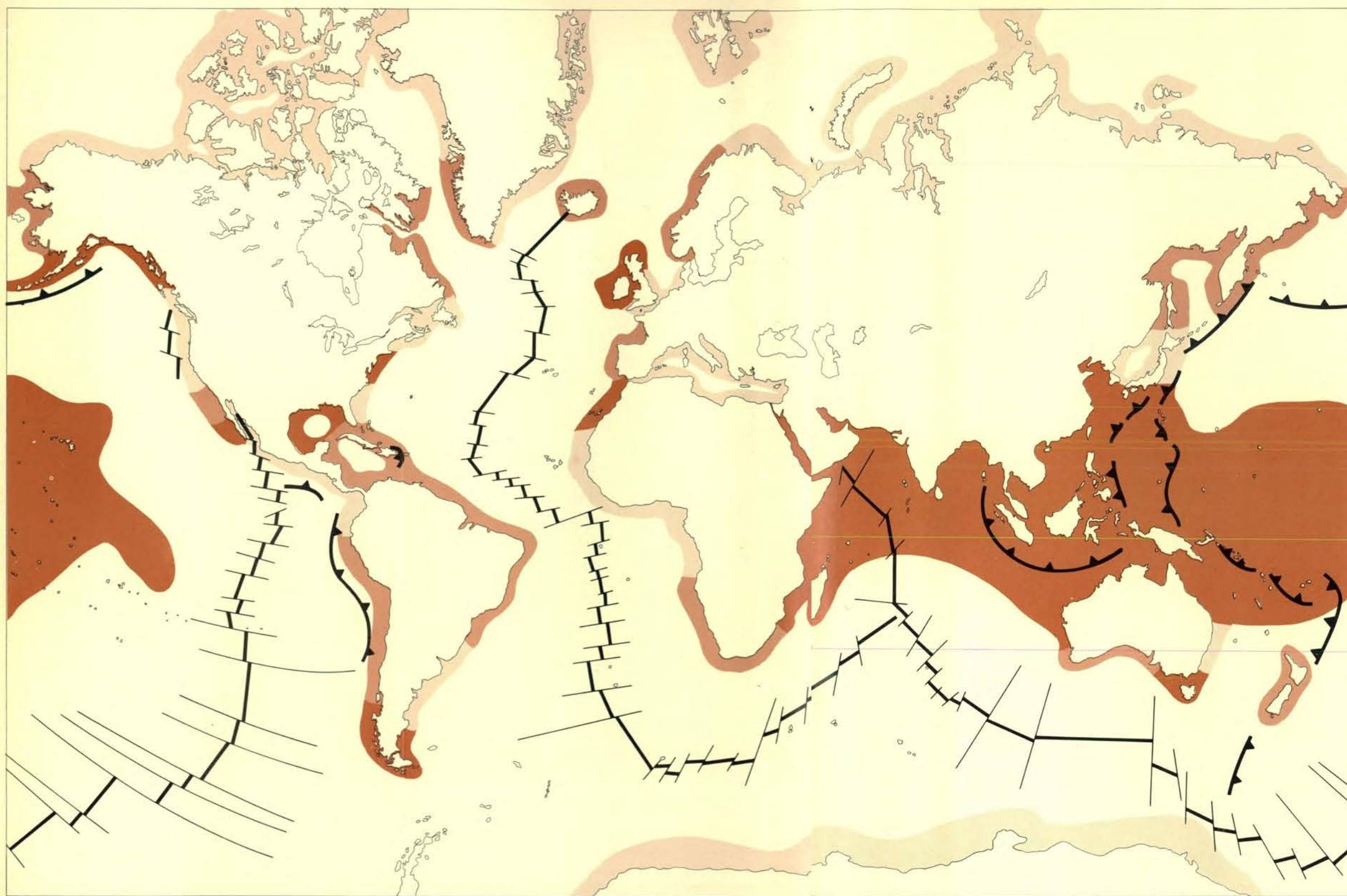
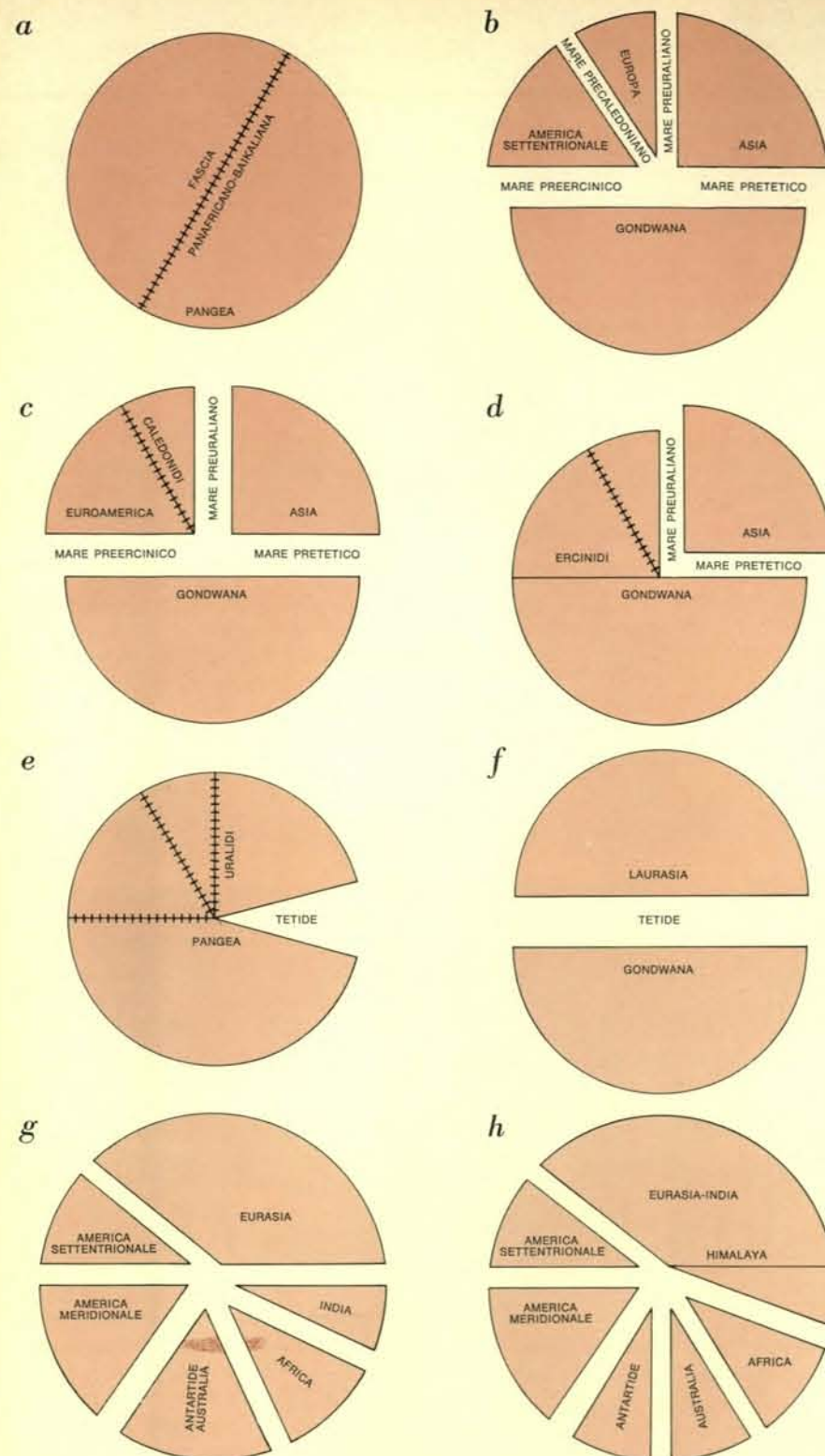
Dato che oggi l'andamento della provincialità e della diversificazione delle faune marine coincide strettamente con l'attuale geografia oceanica e continentale e con la distribuzione conseguente degli ambienti, è lecito aspettarsi che anche la distribuzione delle antiche province e l'andamento della diversificazione nel passato siano stati strettamente connessi con i lineamenti paleogeografici. Una delle forme paleogeografiche meglio conosciute è quella che caratterizzava la Terra all'inizio del Triassico, circa 225 milioni di anni fa. Allora i continenti erano riuniti in un unico supercontinente chiamato Pangea che deve aver avuto un margine continuo di acque basse che lo circondava completamente, senza che vi fossero barriere fisiche importanti a impedire la diffusione degli animali marini di scarsa profondità (*si veda la figura a pagina 59*). Tuttavia, la provincialità deve essere stata bassa se paragonata con quella odierna e deve essere stata causata esclusivamente dal clima. È probabile che il clima marino fosse abbastanza mite e che persino nelle acque delle alte latitudini le temperature fossero molto piú elevate di quanto non lo siano oggi. Ne deriva che la provincialità climatica deve essere stata poco pronunciata. Per di piú i mari, a quel tempo, erano per la maggior parte limitati ai bacini oceanici e non si estendevano in modo significativo al di sopra delle piattaforme continentali. Dunque, l'habitat degli organismi marini d'acqua bassa era notevolmente ridotto, in primo luogo a causa della diminuzione della lunghezza delle linee di costa che viene sempre a determinarsi in seguito alla creazione di un supercontinente da diversi continenti piú piccoli, e in secondo luogo a causa della generale assenza dei mari dalle piattaforme continentali. La riduzione dell'area abitabile gioca a favore di una minore diversificazione delle faune. Da ultimo, l'emergenza di tale supercontinente avrebbe determinato condizioni instabili

Diversità relativa tra le specie bentoniche di scarsa profondità negli oceani attuali. Le classi di diversificazione non sono basate su un conteggio assoluto ma sono dedotte dalla diversificazione dei gruppi dotati di scheletro e meglio noti, come i lamellibranchi, i gasteropodi, gli echinidi e i coralli. Le classi piú elevate (*colore piú scuro*) sono circa 20 volte piú diversificate di quelle inferiori (*colore piú chiaro*).

presso le linee di costa con il risultato di provocare una grande instabilità nella produzione di sostanze alimentari. Tutti questi fattori convergono verso un unico risultato, quello di ridurre la diversificazione delle specie: dunque, sarebbe lecito aspettarsi che le associazioni triassiche avessero larga diffusione e fossero costituite da un numero di specie comparativamente abbastanza ridotto. E questo è esattamente ciò che i resti fossili stanno a indicare.

Prima del Triassico, nel corso del Paleozoico superiore, il grado di diversificazione delle faune sembra essere stato molto più elevato (*si veda la figura a pagina 67 in alto*). Esso si è nettamente ridotto attorno alla fine del Permiano, nel corso di un periodo caratterizzato dall'estinzione di molte specie, che, tutto sommato, sembra essere il piú grave incontrato dalle faune marine in tutta la loro storia. Le specie del Paleozoico superiore, che erano le piú specializzate, si sono quasi completamente estinte, mentre quelle che sono sopravvissute si sono orientate ad avere scheletri semplici. Una parte notevole di questi superstiti sembra essere stata costituita da animali che si nutrivano di detriti sul fondo o in sospensione nell'acqua a contatto con esso. Questi tipi meglio adatti sembrano essere ecologicamente simili alle popolazioni che si trovano oggi negli ambienti instabili, per esempio alle latitudini elevate; i tipi specializzati, invece, sembrano dal punto di vista ecologico, simili alle popolazioni che si trovano negli ambienti stabili, per esempio ai tropici. Cosí le estinzioni sembrano essere state causate dalla riduzione del potenziale di diversificazione per le acque poco profonde, una tendenza associata con una minore provincialità, una ridotta superficie abitabile e meno stabili condizioni ambientali.

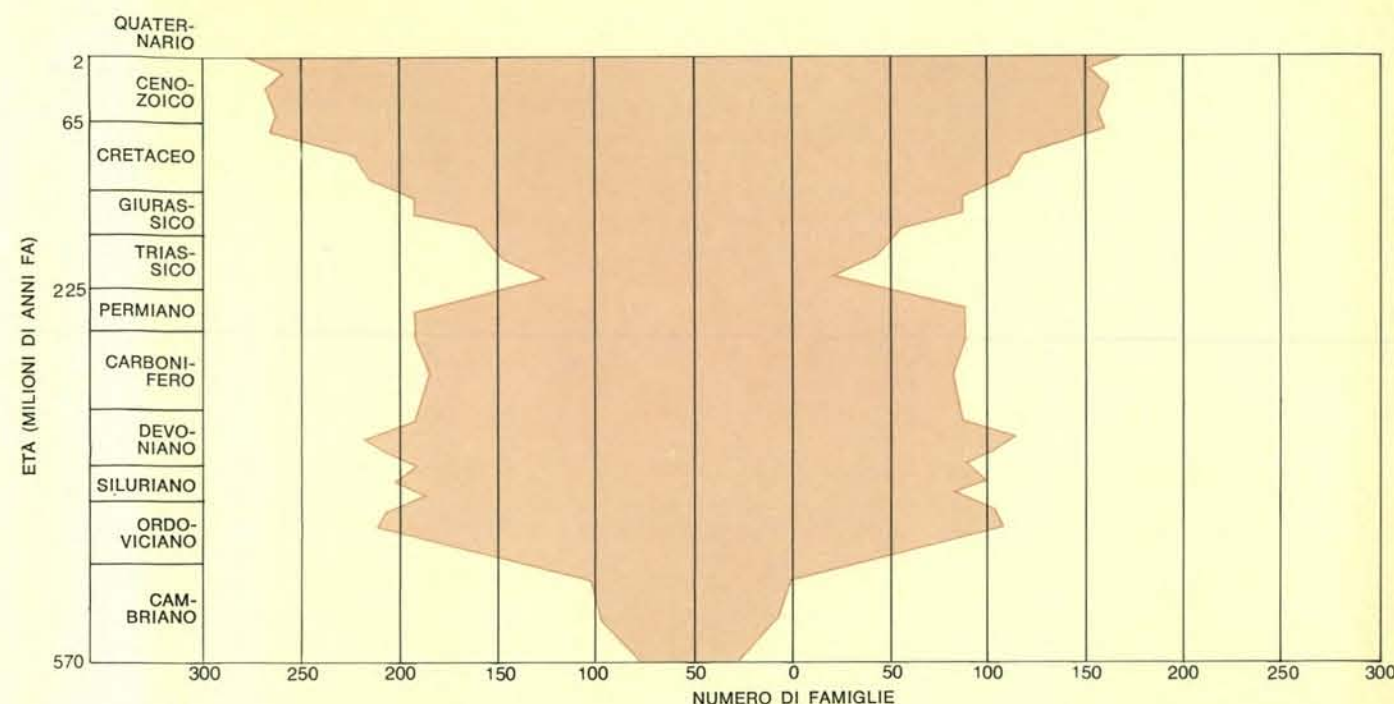
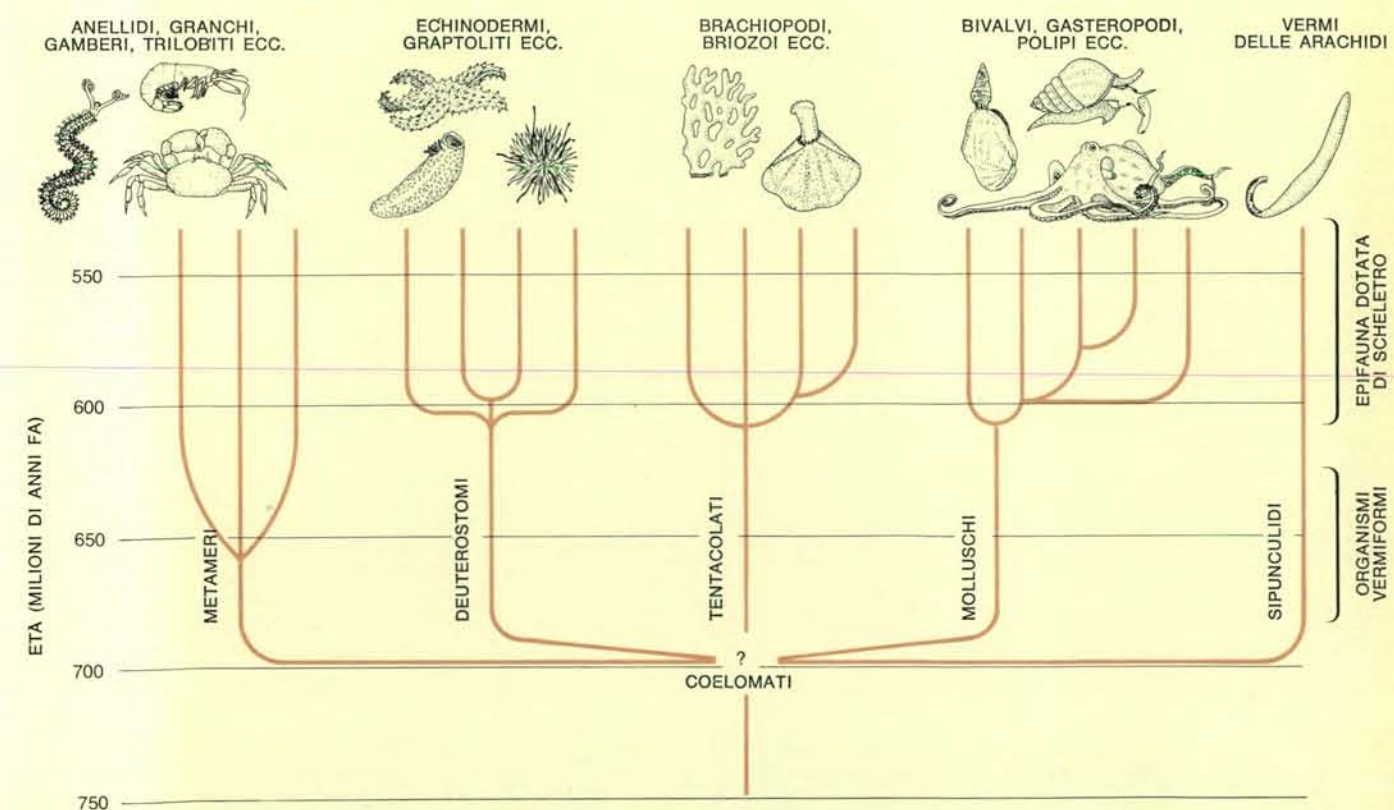
Nel periodo seguito alla grande estinzione, mentre il supercontinente di Pangea si spezzava e i frammenti continentali risultanti migravano fino a raggiungere le attuali posizioni, il grado di provincialità è aumentato, le comunità nelle regioni stabilizzate si sono arricchite di numerosi animali specializzati e la diversificazione globale delle specie nell'oceano mondiale ha raggiunto punte senza precedenti anche se occasionalmente ondate di estinzione hanno interrotto questa tendenza a lunga scadenza.

Accanto al Triassico inferiore vi è, nel passato, un altro periodo in cui bassa provincialità e scarsa diversificazione si sono incontrate con la presenza di un'elevata proporzione di specie che si nutriva di detriti sul fondo o in sospensione presso di esso. Ciò è accaduto nel Precambriano superiore e nel Cambriano, quando una fauna a corpo molle, ad elevata diffusione e a basso grado di diversificazione ha lasciato il passo a una fauna dotata di scheletro, di un minor grado di provincialità e di un grado di diversificazione un poco piú elevato. Sembra probabile che l'ambiente del Precambriano superiore fosse abbastanza instabile e che esistesse o un supercontinente oppure un raggruppamento piú compatto che non oggi delle diverse masse continentali. Nel Precambriano superiore si trovano le prime tracce indiscutibili di invertebrati, ivi comprese forme limicole costituite probabilmente da vermi, dotati di cavità celomatica. Nel Cambriano devono essere esistiti quattro continenti anche se essi non erano distribuiti come oggi. Durante il Cambriano è apparsa una fauna dotata di scheletro che è stata in un primo tempo quasi interamente bentonica e che comprendeva soprattutto forme che si nutrivano di detriti di fondo e in sospensione e, probabilmente, qualche forma erbivora.

Sembra possibile, tuttavia, che le specie del Precambriano superiore fosse-

Le zone in colore indicano le principali province marine di bassa profondità attualmente esistenti. Le catene di province lungo le linee di costa dei continenti derivano dall'elevato gradiente latitudinale nella temperatura oceanica e dalla presenza di barriere sottomarine costituite dalle dorsali di espansione. La grande provincia indopacifica (*colore piú scuro*) ha origine dalla serie di isole sparse.

Schemi semplificati che indicano la configurazione relativa dei continenti e degli oceani nel corso degli ultimi 700 milioni di anni. Il supercontinente del Precambriano superiore (*a*), che esisteva probabilmente circa 700 milioni di anni fa, può essersi formato da due continenti in precedenza separati. Il mondo cambriano (*b*), di circa 570 milioni di anni fa, era costituito invece di quattro continenti. Nel periodo devoniano (*c*), cioè circa 390 milioni di anni fa, vi erano tre continenti in seguito al collasso dell'oceano precaledoniano e alla collisione tra le antiche Europa ed America settentrionale. Nel Carbonifero superiore (*d*), cioè circa 300 milioni di anni fa, Europa e America si sono saldate al Gondwana lungo la fascia ercinica. Nel Permiano superiore (*e*), cioè circa 225 milioni di anni fa, l'Asia si è saldata ai continenti restanti lungo la fascia degli Urali, per formare il continente di Pangea. All'inizio del Mesozoico (*f*), cioè circa 190 milioni di anni fa, Laurasia e Gondwana erano più o meno separati. Verso la fine del Cretaceo (*g*), cioè circa 70 milioni di anni fa, il Gondwana era ormai ridotto in numerosi frammenti e così accadeva in parte anche a Laurasia. L'attuale configurazione (*h*) indica che nel frattempo l'India è giunta a saldarsi all'Eurasia.

ro adattate a condizioni notevolmente instabili e che si siano diversificate soprattutto come associazioni bentoniche costituite da forme che si nutrivano di detriti. La cavità celomatica, evidentemente un adattamento primitivo dei limicoli, si è sviluppata e differenziata in tutta una varietà di forme, fino a costituire forse le cinque forme base: quella dei vermi altamente segmentati che vivevano al disotto della superficie del fondo oceanico e si nutrivano di detriti; quella dei vermi poco segmentati che vivevano fissati al fondo oceanico e si nutrivano di particelle in sospensione; quella dei vermi poco segmentati che vivevano fissati al fondo oceanico e si nutrivano di detriti di fondo; quella dei vermi pseudosegmentati che vivevano sul fondo dell'oceano e si nutrivano di detriti o erano erbivori; e quella dei vermi non segmentati che vivevano al di sotto del fondo dell'oceano e si nutrivano per mezzo di una specie di proboscide. Oltre a questi celomati vi era un certo numero di gruppi di celenterati (come i coralli, le attinie e le meduse) e probabilmente anche vermi piatti e altri vermi non celomati.

Da questi gruppi si sono sviluppate forme di vita animale più elevate; la maggior parte di esse compare nel Cambriano quando evidentemente si sono per la prima volta organizzate nei gruppi che le caratterizzano al giorno d'oggi. Fra i resti fossili di quel tempo appaiono anche animali dotati di scheletro. Probabilmente l'invasione della superficie del fondo marino da parte dei celomati e l'origine di numerose specie dotate di scheletro hanno accompagnato un miglioramento generale delle condizioni ambientali nel momento in cui i continenti si sono dispersi; gli scheletri stessi possono essere visti come forme di adattamento richieste ai vermi per giungere a forme più adatte a vivere sulla superficie del fondo oceanico piuttosto che nel suo interno. L'improvvisa comparsa di scheletri fra i resti fossili, dunque, è associata a una elaborazione generale dei componenti dell'ecosistema marino che vivono sul fondo. Più tardi, dagli abitanti del fondo marino dotati di scheletro, si sono sviluppate forme planctoniche e forme che vivono affondate nei sedimenti, con il risultato che la presenza di uno scheletro è diventata una caratteristica comune in tutto l'ambiente marino.

La correlazione dei principali eventi nella storia della vita con i principali mutamenti ambientali deducibili dal modello della tettonica a zolle è certamente stupefacente. Anche se i particolari dell'interpretazione sono ancora provvisori, sembra sicuro che un ulteriore lavoro di ricerca su queste relazioni darà buoni risultati. In effetti, la capacità dei geologi di determinare la geografia dei continenti del passato, dovrebbe fornire per la prima volta una base per ricostruire la sequenza storica delle condizioni ambientali globali. Tale sequenza potrà allora essere confrontata con il succedersi degli organismi così come appare dai resti fossili. Il tentativo di ricostruzione che segue, eseguito su scala amplissima e trascurando i particolari, costituisce una

Fluttuazione nel numero di famiglie e dunque nel livello di diversificazione, di invertebrati dotati di scheletro e viventi sulle piattaforme continentali del mondo durante gli ultimi 570 milioni di anni. Il tempo geologico procede dal basso all'alto.

Il modello filogenetico dell'evoluzione degli organismi marini con cavità celomatica è basato su deduzioni tratte dal loro adattamento. I gruppi del Precambriano superiore erano costituiti principalmente da vermi che hanno dato luogo nel Cambriano a un'epifauna dotata di scheletro. Gli organismi in alto rappresentano i discendenti attuali dei principali gruppi cambriani.

Dopo la pubblicazione del volume di Letture da LE SCIENZE *dedicato alle*

SCIENZE DELLA TERRA

abbiamo pubblicato sullo stesso argomento i seguenti articoli

I FOSSILI PIÙ ANTICHI
di E.S Barghoorn (n. 36)
Trovati in rocce precambriane chiariscono eventi chiave nell'evoluzione

LA PROSPEZIONE GEOCHIMICA
di M. Dall'Aglio (n. 37)
Sostituisce vantaggiosamente altre prospezioni più costose

LA FAGLIA DI SAN ANDREAS
di D.L. Anderson (n. 42)
Una delle più grandi e importanti dislocazioni della crosta terrestre

L'ORIGINE DEL PETROLIO
di U. Colombo (n. 43)
Origine, migrazione e accumulo di una preziosa fonte energetica

GEOSINCLINALI, OROGENESI E ACCRESCIMENTO DEI CONTINENTI
di R.S. Dietz (n. 46)
Grandi depositi di sedimenti diventano montagne e accrescono i continenti

I CAROTAGGI PROFONDI NEL MEDITERRANEO
di F.C. Wezel (n. 47)
Nuova luce sulla storia geologica di questo bacino

LA TETTONICA A ZOLLE CROSTALI
di J.F. Dewey (n. 48)
La superficie terrestre come mosaico di zolle che interagiscono dinamicamente

GLI ELEFANTI NANI DEL QUATERNARIO DELLA SICILIA
di B. Accordi (n. 49)
Una fauna nana rivela appassionanti problemi ecologici

indicazione del tipo di storia che si sta delineando; esso è basato sugli esempi citati e su analoghe considerazioni.

Nel periodo che ha preceduto i 700 milioni di anni fa, si erano sviluppati animali multicellulari bentonici che rassomigliavano in qualche modo agli attuali platelminti o vermi piatti. Non esistono ancora tracce fossili a testimonianza del cammino evolutivo da essi percorso, ma indizi tratti dall'embriologia e dall'anatomia comparata indicano che essi si sono sviluppati da forme liberamente natanti, forse da forme larvali di meduse che a loro volta si sono sviluppate da animali unicellulari.

Circa 700 milioni di anni fa, forse in relazione all'instaurarsi di condizioni ambientali variabili determinate dall'ammassarsi dei continenti, si è sviluppata una vera e propria cavità celomatica che fungeva da scheletro idrostatico nei vermi cilindrici; questo adattamento ha permesso lo sviluppo di fossatori su fondali marini soffici e ha condotto alla diversificazione di un gruppo di forme diverse di vermi all'atto in cui questo nuovo modo di vivere è stato sperimentato. In alcune rocce del Precambriano superiore sono ancora conservati fossatori di questo tipo. Quando più tardi le condizioni ambientali sono divenute più stabili, diversi gruppi di vermi si sono evoluti per adattarsi a condizioni di vita più variate. Per lo più, i mutamenti nella struttura del corpo necessari per adattarsi a tali condizioni di vita hanno richiesto lo sviluppo di uno scheletro. Vi erano evidentemente tre o quattro tipi principali di vermi che sono rappresentati dai loro discendenti odierni dotati di scheletro. Un tipo era altamente segmentato, come i vermi terricoli, e presumibilmente rimaneggiava di continuo i sedimenti alla ricerca di cibo; nel Cambriano questo tipo è rappresentato dai trilobiti e specie connesse. Un secondo tipo era segmentato in due o tre sezioni celomatiche e scavava debolmente nel fondo per farsi la tana per poi filtrare le particelle alimentari sospese nell'acqua marina appena sopra il fondo oceanico; questi si sono poi evoluti in forme come i brachiopodi e i briozoi. Un terzo tipo era costituito da animali striscianti dotati di una serie di organi interni, ma privi di una vera e propria segmentazione; da questi si sono sviluppate le varie classi di molluschi come i gasteropodi, i lamellibranchi e i cefalopodi Probabilmente, il quarto tipo era costituito da fossatori non segmentati che si nutrivano del detrito superficiale e dai quali sono discesi gli attuali vermi sipunculidi. Da questi possono essere discesi anche gli echinodermi (che comprendono i cetrioli di mare e i ricci) e, da ultimo, anche i cordati e l'uomo. Anche se il filo conduttore dello sviluppo di questi gruppi primitivi e poco noti, è ancora incerto, le tappe dell'adattamento si stanno facendo più chiare.

La principale radiazione cambriana delle specie che vivevano entro i sedimenti verso habitat posti sulla superficie del fondo, ha messo le basi delle linee evolutive fondamentali e ha occupato i principali ambienti marini. Successivi episodi evolutivi si sono manifestati a modificare questi animali primitivi in forme più elaborate. Dopo il Cambriano gli animali marini di scarsa profondità sono divenuti più altamente specializzati e si sono differenziati in un numero più alto di specie, suggerendo una tendenza continua verso la stabilizzazione delle risorse. Gli animali che si nutrivano di particelle in sospensione si sono diffusi e hanno occupato anche le parti più alte della colonna d'acqua; anche i predatori così si sono diversificati sempre di più. Questa tendenza sembra aver raggiunto un massimo nel Devoniano, circa 375 milioni di anni fa. La fauna paleozoica caratteristica è stata alla fine spazzata via nel periodo di riduzione drastica nel grado di diversificazione che ha accompagnato la crisi biologica permotriassica. Così, lo sviluppo della fauna paleozoica si è accompagnato a un miglioramento nelle condizioni ambientali e a un aumento nella provincialità, mentre il declino di quella fauna coincide con il ristabilirsi di condizioni dure e instabili e di un diminuito grado di provincialità. La successiva frattura e dispersione dei continenti ha condotto all'attuale biosfera.

Oggi viviamo in una biosfera altamente diversificata che ospita probabilmente tante specie quante mai sono vissute, fra loro associate in una ricca varietà di comunità animali e in un grande numero di province, probabilmente il più alto che sia mai esistito. Il mondo d'oggi è dotato di una biosfera invidiabilmente diversificata e interessante; sarebbe una tragedia se noi stessi turbassimo l'ambiente in modo tale da provocare un ritorno della biosfera a uno stato di bassa diversificazione e alla contemporanea estinzione di un gran numero di specie. Ovviamente, i processi naturali potrebbero alla fine ristabilire il perduto grado di diversificazione: dovremmo però attendere pazientemente almeno una decina di milioni di anni. Abbiamo l'alternativa di cercare di conservare l'ambiente nel suo stato attuale, di preservare cioè la ricchezza e la varietà della natura.

I vortici nelle scie degli aerei

L'entrata in servizio dei jet di tipo "jumbo" ha trasformato il ben noto fenomeno aerodinamico della turbolenza in un grave pericolo potenziale per gli aerei piú piccoli che vengono a trovarsi nella loro scia

di Norman A. Chigier

La prossima volta che assisterete al decollo di un grosso quadrigetto osservatelo con attenzione e potrete probabilmente vedere che gli scarichi dei motori si presentano all'inizio come quattro strisce di fumo nettamente distinte le quali si uniscono poi in due sole scie separate da una distanza all'incirca pari all'apertura alare dell'aereo. Le strisce di fumo rendono visibili i processi aerodinamici in seguito ai quali si forma, nella scia di ogni aeroplano, una coppia di vortici che si distaccano dall'estremità delle ali. La presenza di vortici di uscita può essere visualizzata in un giorno sereno anche grazie alle lunghe scie bianche di condensazione in coda a un aviogetto che vola ad alta quota. In questo caso i vortici vengono resi visibili dal congelamento del vapor d'acqua nello scarico dei reattori sotto forma di cristalli di ghiaccio che si inseriscono successivamente nel sistema dei vortici di scia.

Il fenomeno dei vortici di uscita dall'ala è noto quasi fin dall'origine dell'aerodinamica. In effetti la formazione dei vortici è stata associata da lungo tempo con la portanza delle ali, ma il pericolo potenziale insito nel fenomeno si è palesato solo da pochi anni. L'entrata in servizio dei grossi aviogetti del tipo jumbo, quali il *Boeing 747*, il *Lockheed C-5A* e il *McDonnell Douglas DC-10*, con il contemporaneo aumento del numero di aerei piú piccoli e piú leggeri, ha trasformato il problema della turbolenza dovuta ai vortici di scia in una grave preoccupazione per la Federal Aviation Administration. Negli ultimi anni il numero di disastri attribuibili alle interferenze fra gli apparecchi e le scie di altri aeroplani è aumentato notevolmente.

Benché la grande maggioranza di incidenti di questo genere abbia coinvolto piccoli velivoli, la caduta nel 1972 di un *DC-9*, trovatosi nella scia di un *DC-10*, con la conseguente perdita dell'aereo e dell'equipaggio, ha dimostrato che anche gli aerei di dimensioni medie possono ora andare soggetti a tali disastri. In seguito a questo incidente, il National Transportation Safety Board ha raccomandato alla FAA che la distanza tra gli aerei negli spazi sottoposti a controllo del traffico venga aumentata ad almeno tre minuti di volo ogniqualvolta esistano condizioni meteorologiche favorevoli alla formazione di vortici. Tre minuti di volo equivalgono a circa 11 chilometri.

Cosa si può fare per cercare di risolvere il problema? Gli obiettivi principali dei programmi di ricerca in corso sono stati rivolti ad accertare la natura, l'intensità e la persistenza dei vortici prodotti da vari tipi di aerei; a sviluppare un metodo pratico di riduzione del rischio per gli attuali velivoli provocando la dissoluzione artificiale dei vortici e infine a installare negli aeroporti sistemi di rivelazione e di controllo per evitare che gli aerei entrino nella turbolenza causata dalla scia dei velivoli di maggiori dimensioni.

I vortici vengono studiati presso gli impianti sperimentali della Federal Aviation Administration vicino ad Atlantic City nel New Jersey. La fotografia in alto nella pagina a fronte mostra un aviogetto jumbo *L-1011* della Lockheed in volo a una quota di 60 metri dietro a una torre appositamente attrezzata da cui vengono emessi strati di fumo colorati che si muovono seguendo con un moto abbastanza regolare la direzione del vento. Nella fotografia in basso le scie vorticose di fumo mettono in evidenza il vortice turbolento che si distacca dalla estremità di una delle ali. È visibile al centro della grande configurazione circolare l'esile e lungo nucleo del vortice rotante ad alta velocità, che si espande solo gradualmente nel tempo. Alcuni attimi piú tardi si formerà il vortice dietro l'estremità dell'altra ala e la relativa scia di fumo comincerà di colpo a ruotare nel verso contrario alla prima. Oltre al dispositivo per la visualizzazione del fumo, la torre, alta 42,70 metri, è equipaggiata con sensori della velocità dell'aria montati ogni 30 centimetri, i cui dati vengono registrati su nastro magnetico.

I modelli fisici dei campi aerodinamici di flusso intorno a un velivolo sono stati determinati sulla base di esperienze di visualizzazione della vena fluida, di previsioni teoriche e di misure dirette della velocità e della pressione sulle superfici dell'aereo. L'aria che si avvicina a un aereo in volo viene deviata verso l'alto e la conseguente deflessione del flusso dà luogo a una forza verticale che fornisce la portanza all'aeroplano. Alle estremità delle ali si forma uno strato limite che si distacca dalla loro superficie e si avvolge in vortici a spirale che hanno origine sul dorso delle ali stesse. I filetti fluidi che si distaccano dalla superficie delle ali convergono in modo che, a una certa distanza dal bordo d'uscita, la configurazione predominante del flusso è una coppia di vortici (*si veda l'illustrazione a pagina 76*), ciascuno dei quali ha aspetto simile ai vortici che si formano quando si apre lo scarico di una vasca da bagno. In entrambi i casi il vortice è costituito da un nucleo centrale che ruota trascinando il fluido che lo circonda. Le velocità periferiche massime si riscontrano in corrispondenza del margine esterno del nucleo, poi diminuiscono esponenzialmente con la distanza dall'asse. L'intensità di un vortice di uscita è direttamente proporzionale al peso del velivolo; di conseguenza, poiché il peso dei piú grossi aerei attualmente in servizio è dell'ordine delle 500 tonnellate, l'energia contenuta in questi vortici può essere notevole.

È possibile calcolare l'energia posseduta da un vortice di scia sulla base delle teorie aerodinamiche correnti. Purtroppo queste teorie, prevalentemente bidimensionali e che di solito

Il dissolvimento brusco di una coppia di vortici di scia dietro un aeromobile è probabilmente causato dall'accrescimento di piccole instabilità idrodinamiche nei nuclei dei vortici. Il numero in alto a sinistra di ogni fotografia indica da quanti secondi è iniziato il vortice. Le fotografie sono state scattate nel corso di un recente programma di prove eseguito dalla FAA per studiare la persistenza e le caratteristiche dei vortici in funzione della configurazione, velocità di volo e quota degli aeroplani. Durante la particolare prova qui illustrata il velivolo, un *Convair 880*, volava a una quota di circa 1500 metri e con una velocità relativa di 300 chilometri all'ora. I vortici vengono evidenziati con olio vaporizzato aggiunto ai gas di scarico dei due reattori esterni. Sequenze fotografiche di questo tipo hanno suggerito l'ipotesi che, se fosse possibile indurre artificialmente delle instabilità nelle scie dei grossi aerei, si potrebbe accelerare la disgregazione della pericolosa struttura dei vortici.

non considerano gli effetti della viscosità, non consentono di prevedere la lunghezza e la persistenza dei vortici né la struttura della zona del nucleo. Recentemente le misure dei campi di flusso nelle scie degli aerei sono state eseguite in gallerie del vento, in vasche d'acqua e in volo. La caratteristica piú preoccupante delle ricerche in volo è stata la constatazione che i vortici sono molto persistenti, modificando la loro intensità solo di poco e a distanze che possono raggiungere perfino i 24 chilometri dalla coda dell'aereo!

Le misure sulla distribuzione delle velocità hanno evidenziato anche la presenza di una componente assiale della velocità nella zona del nucleo, componente che è stata dimostrata in condizioni reali lanciando da terra un certo numero di palloncini gonfiati con elio mentre un velivolo sorvolava il punto della prova. I palloni di elio, sollevatisi sotto l'azione della spinta aerostatica, hanno cominciato a roteare, una volta presi nella scia dell'aereo. Entrati nel campo centrifugo di flusso i palloncini vennero spinti verso l'asse centrale del vortice a causa della minore densità dell'elio rispetto all'aria; raggiunto l'asse i palloncini furono visti spostarsi con una velocità pari a circa un decimo della velocità dell'aeroplano. Velocità assiali dello stesso ordine di grandezza sono state misurate anche in gallerie del vento e in normali condizioni di volo. Si ritiene che questi flussi assiali svolgano una parte importante nel mantenere molto stabili le condizioni all'interno del vortice. I vortici non conservano una posizione fissa nello spazio, ma piuttosto serpeggiano sotto l'azione della turbolenza atmosferica. Questi spostamenti tortuosi hanno costituito una grossa difficoltà per misurare l'energia contenuta nei vortici.

L'autore con i suoi colleghi ha tentato di valutare l'intensità dei vortici di scia in una galleria del vento di 12 × 24 metri mediante un sensore della velocità dell'aria a scansione rapida, chiamato anemometro a filo caldo, presso l'Ames Research Center della National Aeronautics and Space Administration. La sonda, che può registrare le componenti tridimensionali della velocità dell'aria, è costituita in sostanza da tre fili riscaldati elettricamente, ciascuno lungo cinque micron, montati all'estremità di un braccio rotante avente un raggio di 2,44 metri. Il braccio rotante veniva fatto passare nel centro del vortice, mentre i segnali provenienti dai fili caldi venivano registrati su nastro magnetico. Era possibile misurare le componenti della velocità all'interno del vortice calibrando i segnali elettrici forniti dai fili caldi in funzione della velocità, direzione e temperatura del vento. Si potevano cosí determinare le variazioni dell'intensità del vortice lungo la corrente ripetendo l'operazione in corrispondenza di varie stazioni assiali a valle dei modelli fissi di aerei installati nella galleria del vento.

Recentemente i ricercatori dell'Ames hanno adattato un anemometro a filo caldo al longherone di un piccolo aviogetto sperimentale, con il quale sono state eseguite misure sui campi di velocità nelle scie di aerei di maggiori dimensioni in normali condizioni di volo. Un anemometro a laser a scansione rapida, anch'esso costruito all'Ames Research Center, consente di effettuare misure bidimensionali per via ottica sui componenti della velocità in gallerie del vento o in vasche d'acqua senza alterare il flusso. Queste misure hanno dimostrato che i vortici strettamente delimitati sono molto stabili e che cambiano di struttura con lentezza estrema. I principali cambiamenti che si verificano sono un aumento nel diametro del nucleo accompagnato da una riduzione delle velocità periferiche. I vortici distano approssimativamente quanto l'apertura alare dell'aereo e presentano un diametro del nucleo uguale al 3 per cento circa dell'apertura alare. Le velocità periferiche massime misurate sono dell'ordine del 30 per cento della velocità di volo.

È ormai consueto che chi controlla il traffico aereo dia agli aerei in fase di avvicinamento agli aeroporti l'avviso: « Pericolo – turbolenza di scia ». In molti casi il pilota che riceve il messaggio non è in grado di scorgere alcun indizio rilevante la presenza del velivolo che ha generato la turbolenza; cercherà comunque di seguire le istruzioni operative sulla turbolenza di scia, ma sarà ostacolato dal non conoscere l'ubicazione del vortice.

Da una dettagliata analisi di un buon numero di incidenti con perdite di vite e di aerei, sembra chiarito il meccanismo di interazione fra turbolenze e aeroplani. Si può considerare che il campo di flusso nella scia vorticosa agisca come una spinta ascensionale indotta nella zona prossima ai bordi esterni e come una spinta discensionale indotta nella zona compresa fra la coppia di vortici. Un apparecchio che vola lungo la medesima traiettoria seguita dal velivolo generatore, entrando direttamente nel nucleo di uno dei vortici è sottoposto a un improvviso e violento rollio. In tale situazione un aeroplano molto leggero può subire gravi danni strutturali con la conseguente perdita del controllo da parte del pilota. In fase di decollo un aereo che entri nella zona di pressioni discensionali indotte esistente fra i vortici può essere spinto verso il basso.

Il terzo possibile tipo di mutua interferenza si verifica quando l'aereo che segue il generatore della turbolenza attraversa la traiettoria del vortice ad angolo retto. Il pilota avverte un'improvvisa spinta verso l'alto e, quando ha agito sui comandi per bilanciarne l'azione, entra nella zona di pressioni discensionali, che possono spingerlo verso terra.

Una misura quantitativa del pericolo per un aereo coinvolto in una zona di turbolenza vorticosa può essere espressa in funzione dell'energia posseduta dal vortice di uscita. In base alla distribuzione calcolata o misurata della velocità periferica è possibile calcolare l'effettivo momento di rollio del vortice agente sulle ali di un velivolo che per vari motivi entra nella turbolenza di una scia. Sono stati eseguiti studi in gallerie del vento e in vasche d'acqua su tali momenti di rollio misurati direttamente su modelli di piccoli aerei mantenuti nella scia di modelli di aerei di maggiori dimensioni. Gli studi convalidano le valutazioni di elevata pericolosità in cui viene a trovarsi il velivolo che si imbatte nella scia di un altro aeroplano.

Nei velivoli supersonici con ali a delta i vortici si formano al bordo di attacco dell'ala e aumentano di intensità lungo il dorso dell'ala stessa, contribuendo in modo notevole alla portanza.

In condizioni normali di volo una coppia di vortici di scia di solito dapprima perde quota e poi si mantiene orizzontale a una certa altezza. Ciò però non si verifica sempre e vi sono esempi di vortici che si mantengono alla stessa altitudine o che addirittura guadagnano quota su distanze di parecchi chilometri. L'esame delle scie di condensazione indica che alla fine il vortice si disintegra. Le fotografie di queste scie rivelano la presenza di zone di flusso molto stabile e con bassissima dissipazione, seguite da un dissolvimento relativamente brusco. Sono state postulate molte teorie per la spiegazione del fenomeno peraltro non sostenute da verifiche sperimentali.

Per il momento sono stati identificati quattro modi principali di dissolvimento dei vortici: 1) disgregazione, 2) instabilità, 3) dissipazione per viscosità e 4) mutua interferenza con l'atmosfera. Si ha disgregazione quando il nucleo si espande all'improvviso: nel vortice si verifica una rapida transizione da un movimento con momento della quantità di moto fortemente concentrato a un movimento nel quale il raggio del

nucleo aumenta di un ordine di grandezza, con una corrispondente piú ampia distribuzione della quantità di moto. Nella zona di disgregazione si instaurano forti gradienti di pressione, i quali si risolvono spesso in un'inversione del flusso lungo l'asse del nucleo, con un comportamento che si ritiene analogo a quello di un'onda d'urto. La disgregazione dei vortici è stata osservata sia in fotografie di un colorante disperso nella scia di un modello di aereo trainato in una vasca d'acqua, sia su superfici alari a delta con elevati angoli di incidenza.

L'autore ha misurato, nel corso di recenti studi effettuati nella galleria del vento all'Ames, la distribuzione delle velocità all'interno di un vortice, in fase di disgregazione, nella scia di un'ala a delta. Le misure, eseguite con l'ausilio di un laser, hanno mostrato una sensibile riduzione delle velocità assiale e periferica nel nucleo accompagnata dalla generazione di turbolenza. Si sta cercando un sistema per provocare la disgregazione dei vortici nella scia dei grossi aviogetti da trasporto installando sulle ali superfici fortemente inclinate.

In condizioni di elevata stabilità si è osservato che le piccole instabilità si espandono man mano che cresce la distanza dall'aereo generatore. Si tratta di instabilità «idrodinamiche» che provocano un ampliamento del nucleo con conseguente mutua interferenza fra i vortici quando questi si avvicinano l'uno all'altro. Le fotografie prese in volo con fumi traccianti mostrano la congiunzione dei due vortici e la formazione di complicate figure ad anello.

Nei vortici di uscita è sempre presente, in una certa misura, una lieve turbolenza che agisce sui gradienti di velocità producendo sollecitazioni locali. Il flusso rotazionale del vortice ha però un effetto stabilizzante che impedisce la diffusione e la propagazione della turbolenza provocandone perfino lo smorzamento.

I vortici presentano la massima stabilità sopra gli oceani in condizioni atmosferiche calme, mentre nell'atmosfera turbolenta sopra le regioni montuose si dissolvono molto piú rapidamente. Le vaste perturbazioni atmosferiche agiscono sui vortici diffondendo al loro interno la turbolenza dell'aria e provocandone la disintegrazione.

Nella scia di ciascun aeroplano si formano delle coppie di vortici, in seguito a un processo aerodinamico direttamente associato con la portanza generata dalle ali. L'aria deviata verso l'alto dal bordo d'attacco di ciascuna ala forma uno strato limite che si distacca dall'ala stessa e si avvolge in un vortice a spirale (*profili in colore*). A una breve distanza dal bordo di uscita le linee di flusso dello strato limite convergono con il risultato che (indipendentemente dal numero e dalla disposizione dei motori) la configurazione dominante del flusso dietro ciascun aereo è una coppia di vortici separati da una distanza approssimativamente uguale all'apertura alare; il diametro del nucleo vorticoso è pari all'incirca al 3 per cento dell'apertura alare. Si tratta di vortici notevolmente stabili: ne sono stati osservati alcuni che si sono conservati, con pochi cambiamenti, fino a 24 chilometri dietro i grossi aviogetti del tipo jumbo, quali il *Boeing 747* qui rappresentato. Dal punto di vista dell'aereo che segue, il campo di flusso nella scia della coppia di vortici agisce come una spinta ascensionale nella zona prossima ai bordi esterni e come una spinta discensionale nella zona compresa fra la coppia di vortici (*frecce in nero a destra*).

I vortici ad alta quota sono resi visibili dalla formazione di cristalli di ghiaccio nel vapor d'acqua contenuto nei gas di scarico. I cristalli di ghiaccio vengono poi immessi nel sistema vorticoso della scia, con la generazione delle ormai familiari scie di condensazione bianche visibili nei giorni limpidi. L'aereo generatore in questa prova della FAA era un trasporto *C-141* della Aeronautica statunitense. La fotografia tratta da un film a 16 millimetri, è stata ripresa da un velivolo che seguiva il precedente.

La NASA ha avviato nel 1970 un programma fondamentale di ricerche per stabilire se sia possibile attenuare il pericolo dei vortici. Un anno piú tardi la FAA ha richiesto alla NASA di accelerare gli studi. Durante le prime fasi del programma di ricerche i pareri erano divisi fra gli specialisti di aerodinamica teorica e le poche persone che ritenevano fosse possibile trovare qualche metodo per ridurre il pericolo in questione. In base ai principi generali dell'aerodinamica, si sapeva che la presenza e la formazione dei vortici era fondamentalmente connessa con la portanza e, poiché un aeroplano avrebbe inevitabilmente provocato vortici di forte intensità, gli specialisti di aerodinamica che lavoravano nel programma conclusero dichiarando il problema insolubile.

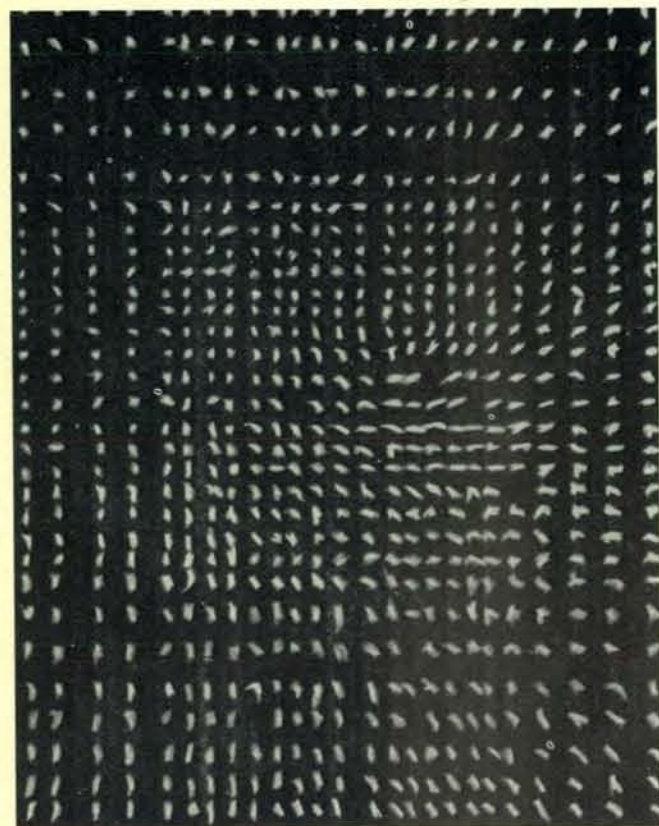
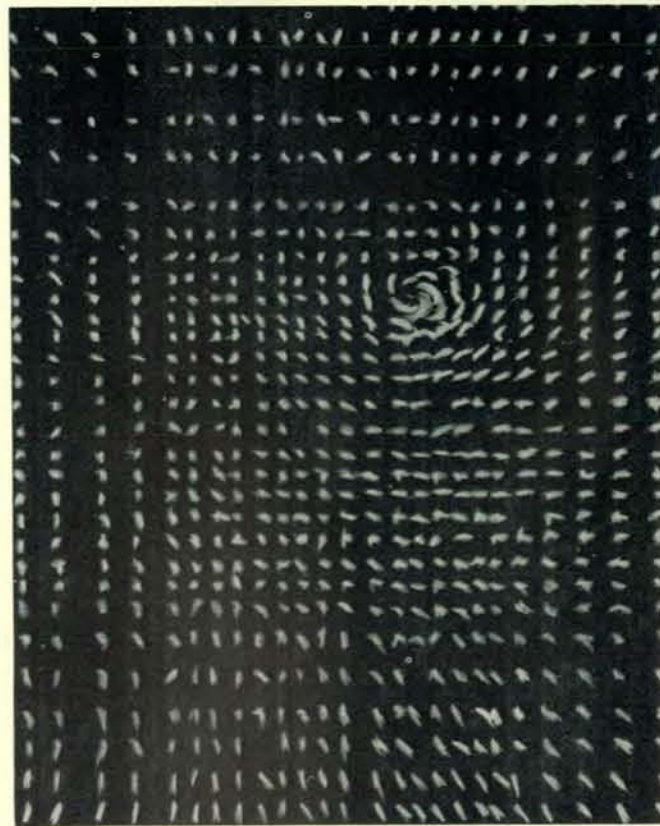
Gli studi preliminari eseguiti in gallerie del vento indicarono però che la presenza di un grosso ostacolo nella zona vorticosa dietro un aereo disgregava la struttura concentrata di un vortice, senza avere alcun effetto sulla portanza o sulla resistenza aerodinamica misurata su un modello di ala. Venne cosí dimostrato che la disgregazione della forma strutturale regolare del vortice non avrebbe interessato le configurazioni del flusso a monte dell'ostacolo, deducendo altresí che ubicare con precisione la zona di disintegrazione del vortice in condizioni atmosferiche normali non avrebbe avuto alcun effetto sulle prestazioni del velivolo. In pratica vi è una piccolissima perdita di momento della quantità di moto di ciascuna coppia di vortici. Tuttavia il momento netto della quantità di moto della scia è nullo cosicché la possibilità che due vortici interferiscano fra loro diffondendo la «vorticosità» in un ampio spazio non viola le leggi fondamentali dell'aerodinamica. Da un punto di vista pratico, se la vorticosità si distribuisce in una zona assai piú ampia, le velocità diventano molto basse e le forze esercitate su un aereo che segue risultano trascurabili.

Il programma di ricerche della NASA si concentrò quindi sui mezzi pratici per accelerare il dissolvimento dei vortici. Il primo metodo sperimentato era piuttosto «brutale» e si basava sulla formazione di una turbolenza all'interno del vortice sufficiente a disgregare il sistema di flusso stabile, regolare e simmetrico rispetto all'asse. Le prove in galleria del vento eseguite all'Ames Research Center nel 1970 dimostrarono che un «diruttore» montato verticalmente sul dorso dell'ala vicino all'estremità, dove si forma il vortice, era in grado di ridurre le velocità periferiche massime del 63 per cento con un corrispondente aumento del diametro del nucleo. Il diruttore provocava anche una diminuzione di portanza e un aumento di resistenza.

Venne quindi adattato un diruttore all'ala di un aviogetto da trasporto *Convair 990*. Un aereo per ricerche di minori dimensioni (un jet *Lear*) fu fatto volare seguendo approssimativamente l'asse del vortice, mentre venivano effettuate misure sulla risposta dinamica del velivolo. Per rendersi conto del disagio subito dai passeggeri, l'autore viaggiava al posto del secondo pilota. Il velivolo venne anzitutto posizionato al di sopra del vortice, con l'aiuto di fumo aggiunto ai gas di scarico del *Convair 990*, e quindi fatto discendere verso il vortice stesso. Lo scuotimento riscontrato fu del tutto diverso da quello subito per l'incontro accidentale con una turbolenza in aria chiara. L'autore ebbe l'impressione che qualcuno avesse agguantato le ali facendo ruotare con violenza il velivolo e, sebbene allacciato al seggiolino con la cintura di sicurezza, urtò con la testa contro il tetto della cabina mentre il pilota perse completamente il controllo dell'aviogetto, che fu espulso dal vortice per entrare pochi secondi dopo in quello adiacente con un analogo rollio di scuotimento in senso contrario al primo. Né il pilota, né l'autore furono in grado di percepire differenze sostanziali fra le penetrazioni nel vortice eseguite a 8, a 4,8 e a 1,6 chilometri di distanza dall'aereo generatore della turbolenza.

Sull'aereo di ricerca vennero registrate le variazioni nella deflessione degli alettoni, nel ritmo del rollio e nell'angolo di scivolata. Le registrazioni dimostrarono che il pilota era in grado di mantenere il controllo per un periodo leggermente piú lungo e con minori sbandamenti in un vortice prodotto da un'ala fornita di diruttore. L'autore, da parte sua, si sentí ugualmente male in entrambi i vortici e ritiene che qualunque fosse l'effetto di dissipazione prodotto, non sia possibile considerarlo accettabile per i normali passeggeri.

Dopo queste prove preliminari si applicarono molti tipi di dispositivi a modelli di aeroplani che vennero poi sottoposti a prove in gallerie del vento e

Un dispositivo sperimentale per il dissolvimento dei vortici, costituito da un piccolo «diruttore» verticale, è stato provato dall'autore e dai suoi colleghi in una galleria del vento all'Ames Research Center della National Aeronautics and Space Administration. L'effetto del diruttore è stato studiato visualizzandolo con fumo e con una griglia munita di nastri (*a destra*).

Le fotografie di una griglia munita di nastri hanno messo in evidenza che l'installazione di un diruttore può, in certe condizioni, ridurre in larga misura la velocità periferica massima di un vortice con un corrispondente aumento del diametro del nucleo. Il diruttore provoca anche una piccola diminuzione di portanza e un notevole incremento nella resistenza aerodinamica. La configurazione del vortice modificato è mostrata a sinistra, e quella del vortice non modificato a destra. In entrambi i casi la distanza fra l'ala e la griglia era pari a 40 volte la corda dell'ala. Le prove in volo dei dispositivi a diruttore non hanno per ora approdato a risultati conclusivi; vi sono inoltre alcuni indizi che il vortice si riformi a valle del diruttore.

in vasche d'acqua. Le soluzioni sperimentate comprendono forti getti soffiati nel nucleo del vortice, ancore galleggianti montate dietro il bordo d'uscita dell'ala, pannelli fissi applicati in varie posizioni sul dorso dell'ala e deflettori fissati con vari angoli lungo il bordo d'uscita. Sono tutti dispositivi che generano turbolenza, la cui convezione e diffusione all'interno del vortice provocano l'apertura del nucleo. Sebbene gli esperimenti siano ancora in corso, vi sono prove che l'uso di deflettori o di alette di estradosso per sgravare le estremità delle ali e per provocare oscillazioni nel carico sull'apertura alare riduca il momento di rollio della scia. Tali dispositivi introducono un aumento nella resistenza aerodinamica, ma di entità tale da non essere necessariamente proibitiva. Le prove in vasche d'acqua dimostrano che quando il vortice viene fatto espandere in prossimità dell'ala, non si attenua con la medesima rapidità del vortice non modificato, cosicché alla distanza di 8 chilometri dietro l'aeroplano le massime velocità indotte per effetto del vortice sono le medesime. Ciò significa che si può ottenere un considerevole miglioramento solo per distanze inferiori a 8 chilometri.

In definitiva i tentativi di inserire delle instabilità nel vortice possono condurre a soluzioni piú promettenti. Deflettori fatti oscillare sinusoidalmente, in modo da non diminuire la portanza netta, determinano l'oscillazione dei vortici con conseguente mutua interferenza e successiva formazione di cappi a distanze minori che nel caso in cui i deflettori non vengano impiegati. Un pilota è riuscito a ottenere un rapido dissolvimento dei vortici facendo oscillare la barra di comando dell'aereo generatore di turbolenza; in tal caso, tuttavia, il disagio è stato semplicemente trasferito ai passeggeri dell'aereo generatore!

Un ulteriore tentativo di affrontare il problema è costituito dalla variazione del carico alare, procedimento che implica il cambiamento della forma dell'ala e di conseguenza la distribuzione della pressione sulle sue superfici. In linea teorica si può dimostrare la possibilità di progettare un'ala nella quale le linee di flusso non si raccolgano in una coppia di vortici concentrati. L'aviogetto *C-5A* è già equipaggiato con alettoni e con ipersostentatori posteriori a segmenti (usati entrambi per motivi strutturali); si tratta di due sistemi pratici per produrre carichi appropriati in grado di mitigare il pericolo dei vortici.

Frattanto i piloti continuano a riferire di frequenti incontri con turbolenze dovute a scie vorticose, mentre crescono i casi di vertenze con la FAA, l'agenzia considerata responsabile negli Stati Uniti di tali scontri in prossimità degli aeroporti. Gli incidenti che hanno causato gravi danni o perdite di vite umane hanno ormai superato il centinaio. Supponendo che l'attenuazione dell'attuale crisi energetica consenta la ripresa dei recenti orientamenti del mercato, la situazione peggiorerà con l'entrata in servizio di nuovi aviogetti del tipo jumbo e con l'acquisto da parte di un maggior numero di privati di aerei leggeri. La consuetudine negli Stati Uniti è di consentire il decollo e l'atterraggio degli aerei sulla base dell'ordine di presentazione, anche se i controllori del traffico aereo informano con maggiore frequenza i piloti dei movimenti locali dei jumbo.

La FAA sta ora concentrando i propri sforzi per rivelare e tenere sotto controllo i vortici in prossimità degli aeroporti. Sono in corso di sviluppo radar Doppler, anemometri a laser e rivelatori acustici per misurare sia l'intensità, sia l'ubicazione dei vortici. Attualmente una delle apparecchiature piú promettenti è un rivelatore radar-acustico, con il quale un segnale acustico trasmesso da terra viene riflesso dal vortice. Il segnale riflesso viene misurato da un ricevitore situato a una certa distanza dal generatore di rumore. Un certo numero di questi apparati è stato installato in via sperimentale nell'aeroporto internazionale O'Hare di Chicago e in altri aeroporti. Però la portata di queste apparecchiature è limitata a poche centinaia di metri e inoltre il rumore che irradiano è molesto.

Sono stati eseguiti dettagliati progetti di sistemi da installare nelle torri di controllo del traffico aereo per effettuare previsioni delle zone pericolose con il calcolatore, in base al traffico previsto, tenendo conto del peso e dell'assetto di volo degli aerei come pure delle condizioni atmosferiche locali. Gli aerei vulnerabili che si avvicinano o si allontanano dall'aeroporto verrebbero dirottati lontano dalle zone pericolose. Questi sistemi sono molto costosi e, per renderli sufficientemente affidabili per consentire agli aerei di volare a distanze ravvicinate, dovranno essere spesi molti altri milioni di dollari per il loro sviluppo. Vi è però fra gli specialisti un indirizzo secondo il quale la probabilità che un velivolo incontri la piccola zona del nucleo vorticoso è tanto remota che sarebbe sufficiente impartire ai piloti una direttiva di volo di «elusione generica». Le eventuali cadute di aerei dovrebbero perciò essere considerate degli sfortunati incidenti. È ovvio che, qualora il numero di disgrazie aumentasse, tale soluzione non sarebbe piú accettabile.

La tendenza dei futuri sviluppi aeronautici è verso gli aviogetti di tipo jumbo ancora piú grossi degli attuali. Sono in corso di progettazione per collegamenti a breve e medio raggio aeroplani STOL (*short takeoff and landing*, decollo e atterraggio corti) e VTOL (*vertical takeoff and landing*, decollo e atterraggio verticali). Le prove di volo preliminari su aerei STOL con «ipersostentatori potenziati» hanno indicato che le velocità dei vortici vengono attenuate molto piú rapidamente di quanto previsto nel caso della circuitazione equivalente per un'ala senza portanza ottenuta mediante i motori. I tipi piú recenti di aerei jumbo, il *DC-10* e il *Lockheed L-1011*, sono equipaggiati con tre motori a reazione, uno dei quali disposto in posizione centrale sopra la fusoliera. Le prove di volo eseguite nella scia di questi trigetti hanno dimostrato che, quando gli ipersostentatori sono estratti, i vortici risultano perfino piú intensi di quelli causati da un quadrigetto di peso equivalente, mettendo cosí in evidenza che anche l'ubicazione dei gas di scarico può avere influenza sulla stabilità dei vortici. Gli aeroplani supersonici in volo nella stratosfera, ove esistono condizioni di alta stabilità, potrebbero produrre scie di lunghezza molto maggiore di quelle provocate a quote minori nella piú turbolenta atmosfera. Se gli aeromobili di tipo STOL e VTOL, ivi compresi gli elicotteri, verranno adottati come principale mezzo di trasporto a corto e medio raggio, la densità del traffico aumenterà in modo considerevole. Malgrado sia stata prestata solo scarsa attenzione alle scie di questi aeromobili, i loro elevati valori di portanza potrebbero causare, almeno in linea di principio, robusti vortici. Dopo il passaggio di un elicottero sono stati osservati dei vortici consistenti che hanno persistito per vari minuti.

Le previsioni sugli sviluppi futuri dei progetti aeronautici e sulla densità del traffico aereo indicano che le attuali situazioni di pericolo tenderanno a peggiorare; gli sforzi in corso per trovare alcuni metodi di riduzione dell'intensità dei vortici dovranno quindi essere intensificati. È già stato suggerito che le enormi quantità di energia racchiusa nei vortici dovrebbero essere sfruttate per l'azionamento di turbine. Come in molti altri campi della tecnologia, appena l'opinione pubblica riterrà indispensabile per la propria sicurezza risolvere questo problema, sarà molto piú facile reperire i fondi necessari.

Il pianto del neonato

Lo spettrografo acustico rivela le nette differenze fra il pianto di bambini normali e quello di bambini anormali. Questa tecnica può essere utilizzata nella scoperta precoce di disturbi mentali

di Peter F. Ostwald e Philip Peltzman

Che cosa rende così irresistibile il pianto di un bambino nei primi mesi di vita? Questo suono stimola intensi sentimenti e reazioni ben precise in quasi tutti coloro che lo odono. Senza dubbio l'effetto del pianto del bambino ha in gran parte una determinazione biologica, la quale ha come conseguenza di garantire che il bambino riceva cure e nutrizione. Il pianto è uno dei primi modi con cui il bambino riesce a comunicare col mondo. Una madre impara a capire assai presto le differenze tra le forme di pianto del bambino, anche se non ne interpreta sempre nel modo giusto il significato. Nella nostra ricerca al Langley Porter Neuropsychiatric Institute di San Francisco abbiamo cercato di sviluppare metodi di analisi più precisi e più obiettivi dei suoni emessi dal bambino nei primi mesi di vita. Con l'ausilio di varie tecniche acustiche siamo riusciti non soltanto a identificare il pianto di dolore di bambini normali, ma anche a scoprire, attraverso il pianto caratteristico dei bambini affetti da disturbi psichici, la presenza di talune anormalità e infermità.

Una tra le prime comunicazioni obiettive sulla struttura acustica delle grida infantili fu pubblicata nel 1832 da William Gardiner. Nel libro *The Music of Nature* egli descriveva per mezzo di note musicali le grida e i richiami di esseri umani e di animali. « I bambini – egli scriveva – non hanno alcuna difficoltà nell'esprimere desideri, piaceri e dolori mediante le loro grida, assai prima di conoscere l'uso o il significato di una parola ». Tra le sue descrizioni vi erano quelle della « voce dispettosa di un bambino che ne molesta capricciosamente un altro » e del « piagnucolio di un bambino viziato ». Secondo Gardiner, i toni delle grida dei bambini sono compresi generalmente fra le note la e mi, nella parte centrale della tastiera del pianoforte. La componente espiratoria è di solito il carattere più cospicuo del grido; essa dura in media un secondo circa e ha una struttura melodica oscillante. La componente inspiratoria del grido è molto più breve.

Anche Charles Darwin si interessò al pianto dei bambini. Nel libro *The Expression of the Emotions in Man and Animals (L'espressione delle emozioni nell'uomo e negli animali)*, edito nel 1872, presentò una serie di fotografie che illustravano le smorfie associate, nei bambini, all'espressione di emozioni attraverso il pianto. Quando Thomas Edison inventò il fonografo, divenne possibile studiare un grido attraverso la sua riproduzione ripetuta e determinarne la struttura. Nel 1906 Theodor S. Flatau e Hermann Gutzmann, in Germania, registrarono su cilindri di cera di Edison il pianto di bambini e ne diedero una descrizione, rimasta classica, delle proprietà acustiche. Per caratterizzare i caratteri sia vocalici sia consonantici del pianto infantile essi usarono sia la notazione musicale sia l'alfabeto fonetico. Fra i bambini le cui voci furono registrate uno aveva disturbi respiratori; questo fatto fu avvertito immediatamente perché l'altezza del suono era superiore di un'ottava rispetto alla norma degli altri bambini.

Per alcuni decenni lo studio dei suoni emessi dai bambini fu proseguito essenzialmente usando il metodo sviluppato da Flatau e da Gutzmann. Non si fecero grandi passi avanti fino allo sviluppo dello spettrografo acustico nei Bell Telephone Laboratories fra il 1940 e il 1950. Questo strumento scompone suoni complessi in bande di frequenze, che presenta in una registrazione orizzontale. Nel 1951 A.W. Lynip, dello Houghton College, nello Stato di New York, dimostrò per mezzo di spettrogrammi acustici che esiste una differenza fra le grida di un bambino che ha fame e quelle tendenti a richiamare l'attenzione. Lo spettrografo acustico divenne ben presto un elemento standard dell'apparecchiatura richiesta nello studio di suoni infantili.

L'analisi spettroacustica delle grida infantili rivela che la frequenza fondamentale del pianto del neonato si colloca normalmente in una banda compresa fra 400 e 600 hertz (cicli al secondo). La risonanza a frequenze superiori tende a raggiungere i suoi valori massimi fra i 1200 e i 4800 hertz. Nel caso di grida in espirazione causate da dolore, il tono fondamentale può salire temporaneamente di varie ottave o spezzarsi bruscamente in un rumore di larga banda. I neonati non emettono naturalmente solo grida di sofferenza; essi possono anche mormorare o produrre un sommesso chiacchiericcio. Tali suoni vengono emessi inizialmente solo quando il bambino è tranquillo o quando sta poppando al seno, ma è probabile che abbiano un ruolo importante più tardi nell'espressione dei sentimenti e nella comunicazione. Il chiacchiericcio è probabilmente il primo suono di approvazione emesso dal bambino. La sua durata e la sua qualità sono molto più vicine a suoni fonetici che non quelle del pianto e delle grida e molti linguisti ritengono che tali suoni rappresentino il vero inizio della parola.

Il fatto che i bambini producano grida tipiche e che imparino rapidamente a coordinare il loro comportamento con quello della madre ha indotto numerosi ricercatori a investigare strutture di comportamento normali e anormali delle grida, del mormorio e di al-

tre vocalizzazioni durante il primo anno di vita. Samuel Karelitz, del Downstate Medical Center dell'Università di Stato di New York, ha compiuto una registrazione fonografica la quale dimostra che le grida di bambini affetti da taluni disturbi cerebrali o da talune anomalie genetiche differiscono dalle grida di bambini normali. Henry M. Truby, dell'Università di Miami, usò spettrogrammi per stabilire relazioni fra i suoni emessi e i movimenti della bocca e della laringe del bambino. Le ricerche di Ole Wasz-Höckert e dei suoi collaboratori in Finlandia e di John Lind in Svezia hanno fornito prove evidenti del fatto che le analisi spettrografiche del pianto infantile hanno un preciso valore per la diagnosi di numerose condizioni infettive, malattie metaboliche e anomalie di sviluppo.

Il nostro laboratorio di registrazione acustica ha sede nella sezione maternità di un ospedale, cosicché possiamo studiare i neonati subito dopo la nascita. Il neonato viene collocato in una camera di registrazione in cui la temperatura è mantenuta fra i 26 e mezzo e i 29 gradi centigradi. In media si registrano 25 grida per ogni bambino. Quando lavoriamo su bambini malati o prematuri che devono essere tenuti nell'incubatrice, questa può essere spostata nella camera di registrazione senza arrecare alcun disturbo al bambino. Per registrare i suoni emessi da un bambino, un microfono molto sensibile è sospeso sopra di lui. I suoni vengono registrati su nastro e dalla registrazione vengono eseguiti spettrogrammi acustici. Le caratteristiche di frequenza e di ampiezza dei suoni possono essere analizzate anche direttamente per mezzo di un calcolatore digitale programmato in vista dell'elaborazione di tali dati.

Abbiamo concentrato la nostra attenzione sulle grida di sofferenza dei neonati. Abbiamo registrato numerosi suoni di questo tipo determinati da operazioni mediche di routine, come il prelevamento di campioni di sangue e interventi di piccola chirurgia, come la circoncisione. Il grido di sofferenza è più forte, più lungo e più rumoroso del pianto determinato dalla fame. Esso tende anche a essere irregolare, con più interruzioni e smorzamenti.

Talune volte un tipo di pianto insolito fornisce la prima indicazione del fatto che nel bambino c'è qualcosa che non va. Un'infermiera notò un pianto stranamente acuto in un bambino che era considerato del tutto sano alla nascita. Non c'era nulla di insolito nel peso, nella respirazione, nel colore o nel comportamento. Seguendo attentamente il bambino, l'infermiera si accorse però che egli presentava oscillazioni molto forti nella temperatura corporea. Un esame compiuto da un neurologo condusse alla scoperta che il bambino aveva un grave difetto congenito: era privo di corteccia cerebrale.

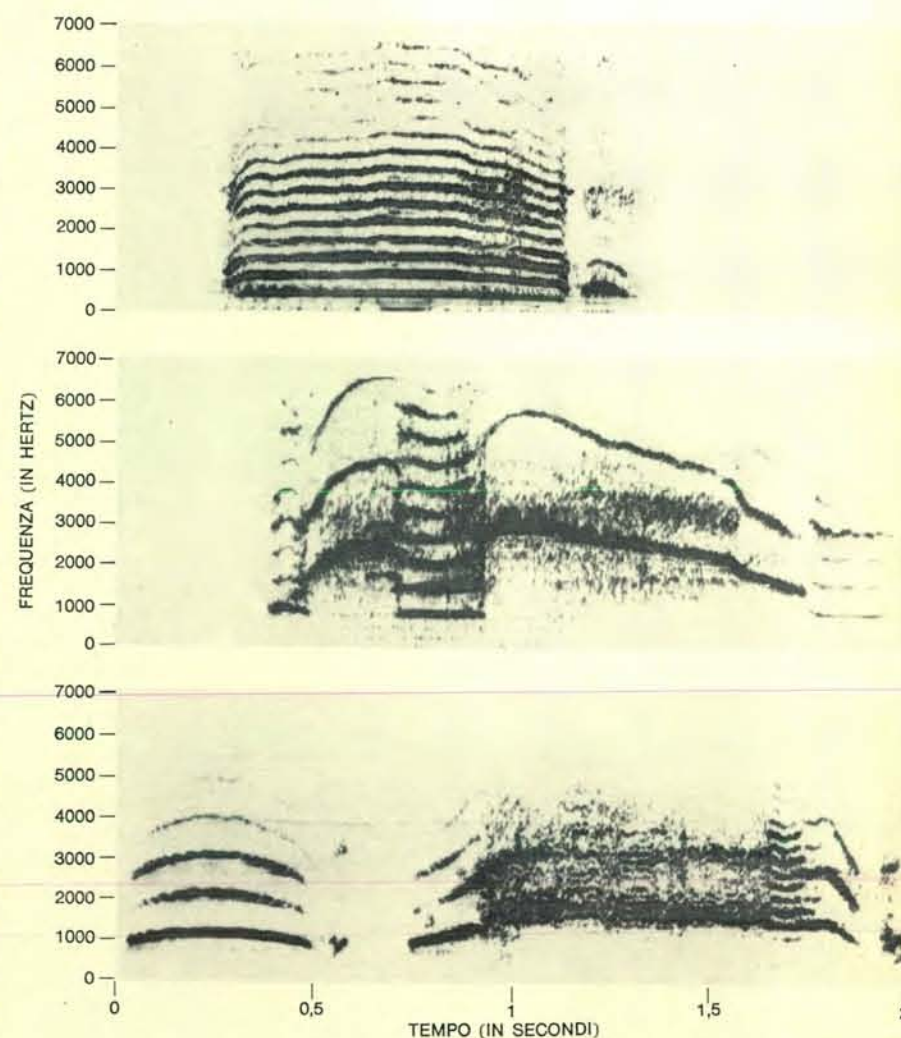
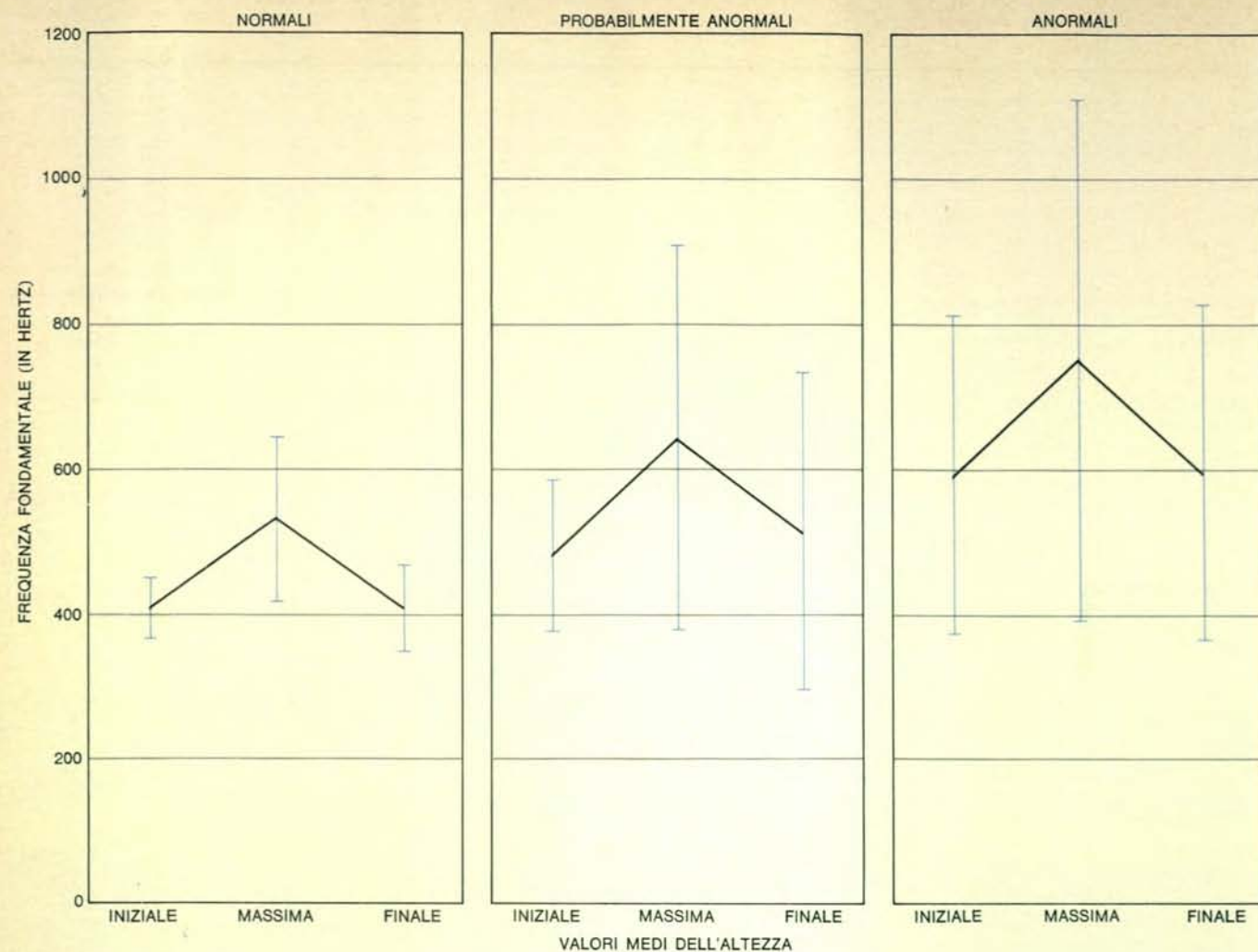
In un altro caso avemmo gravi difficoltà nell'ottenere grida da un neonato di due giorni. L'infermiera ci disse che il bambino piangeva raramente, cosa assai insolita in neonati sani, a meno che non siano sotto l'effetto di sedativi. Quando il bambino gridava, il suono era piuttosto alto. Benché un esame pediatrico a quell'epoca non rivelasse alcun segno clinico di infermità, nelle settimane successive fu chiaro che il bambino non si sviluppava normalmente. Un esame fisico approfondito rivelò un grave difetto nella circolazione sanguigna fra il cuore e i polmoni.

Al fine di accertare se la durata del pianto e l'altezza del grido fossero in qualche relazione con le valutazioni clinico-diagnostiche dello stato di salute dei neonati, registrammo le grida di sofferenza di 13 neonati ospiti nel reparto maternità dell'ospedale. Verificammo le cartelle cliniche relative a ogni neonato, prendendo nota dei particolari della storia prenatale, della nascita e dello sviluppo postnatale. Cinque bambini erano normali; cinque erano probabilmente anormali, benché non fosse stata diagnosticata per essi alcuna infermità specifica; tre erano decisamente anormali. Nella durata del pianto dei bambini dei tre gruppi clinici non c'era alcuna differenza significativa. C'era invece una differenza fondamentale nell'altezza dei suoni emessi dai bambini normali, da quelli probabilmente anormali e da quelli decisamente anormali. Trovammo che i bambini probabilmente anormali e quelli per i quali l'anormalità era stata accertata emettevano suoni di altezza superiore rispetto a quelli emessi da bambini normali (*si veda la figura a pag. 88*).

Un'altra tecnica diagnostica promettente per il rilevamento di anormalità nei neonati, oltre all'analisi spettrografica dei suoni emessi, è la misurazione dei potenziali uditivi evocati, ossia le variazioni nelle onde cerebrali in risposta ai suoni. Elettrodi applicati alla testa del bambino avvertono le piccole variazioni di potenziale elettrico nel cervello che hanno luogo in risposta ai suoni presentati al neonato per mezzo di auricolari in miniatura. Le variazioni elettriche sono poi analizzate da un calcolatore. Possono essere studiati anche i potenziali evocati visivamente se gli stimoli sonori vengono sostituiti da una luce lampeggiante.

L'utilità di questa tecnica elettrofisiologica in combinazione con l'analisi spettografica è stata dimostrata in un caso clinico insolito. Una coppia di gemelle erano nate unite per la testa.

I suoni emessi dai bambini sono registrati in una camera acusticamente isolata. Per mantenere il bambino a suo agio, si fa circolare nella camera aria riscaldata a 26-29 gradi centigradi. I bambini sono collocati di solito su un soffice materasso: nel caso di bambini malati o prematuri l'incubatrice può essere spostata nella camera di registrazione. Un microfono molto sensibile è sospeso a 6-12 centimetri dalla bocca del neonato: l'altezza precisa è stabilita in base all'età e all'energia vocale di ogni bambino. I suoni vengono dapprima registrati su un registratore le cui caratteristiche acustiche siano ben note, e poi vengono analizzati.

La notazione delle grida per mezzo di segni musicali fu introdotta nel 1832 da William Gardiner in *The Music of Nature*, in cui erano analizzate le grida e i richiami di esseri umani e di animali. Ecco alcuni esempi della notazione usata da Gardiner: vagiti di bambino (*a*), pianto di bambino (*b*, *c*) e «piagnucolio di bambino viziato» (*d*).

Spettrogrammi acustici di pianto di neonati sofferenti rivelano la differenza nella struttura acustica delle grida di bambini normali e anormali. Il grido di sofferenza di un neonato normale ha una tipica componente espiratoria prolungata seguita da una breve inspirazione affannosa (*spettrogramma in alto*). Il grido di sofferenza di un bambino diagnosticato come probabilmente anormale era più irregolare e più acuto (*spettrogramma in centro*). Questo bambino presentò più tardi un ritardo generalizzato nello sviluppo motorio e mentale. Un bambino diagnosticato come anormale emetteva grida acute, irregolari, alcune di esse prive di turbolenza e altre con turbolenza (*in basso*).

Le strutture dell'altezza di grida di bambini appartenenti a tre gruppi clinici (neonati normali, probabilmente anormali e sicuramente anormali) presentano nette differenze. Furono analizzate in totale 356 grida di sofferenza, emesse da 13 bambini. Per ogni gruppo clinico furono determinate le altezze medie iniziale, massima e finale. Le linee colorate verticali rappresentano inoltre l'intervallo di variazione attorno alla media. Tali linee indicano che le grida di neonati anormali raggiungono un'altezza di tonalità che va nettamente oltre la gamma dei bambini normali. Queste scoperte suggeriscono la possibilità di sviluppare un test di controllo precoce per la scoperta di anormalità nei bambini attraverso l'analisi delle loro grida.

Non era certo se le due bambine avessero o no ciascuna un sistema nervoso funzionante in modo indipendente. Quando eseguimmo le prime registrazioni dei suoni emessi dalle gemelle in età di tre mesi, le grida della bambina *A* erano piú forti e piú prolungate di quelle della bambina *B*. La bambina *A* gridava anche piú a lungo della gemella, i cui suoni erano deboli e irregolari.

Gli spettrogrammi acustici della bambina *A* avevano esattamente lo stesso aspetto di quelli che siamo pervenuti a considerare tipici di un bambino normale. Il suono saliva in altezza, rimaneva relativamente costante e cadeva poi in modo graduale, il tutto all'interno di un'esalazione di respiro che durava un secondo circa. Gli spettrogrammi acustici della bambina *B* erano invece chiaramente anormali. Spesso durante un'espirazione c'erano interruzioni, cosí brevi da non riuscire percepibili a meno di ascoltare molto attentamente. La bambina *B* emetteva inoltre piú grida corte e piú grida lunghe della bambina *A*. Un confronto fra la intensità totale delle grida rivelò che i suoni emessi dalla bambina *B* avevano un'energia acustica significativamente minore di quelli emessi dalla bambina *A*.

I risultati dell'analisi acustica indicavano che il sistema nervoso della bambina *A* funzionava chiaramente in modo normale mentre quello della bambina *B* presentava probabilmente qualche danno. Sappiamo che la vocalizzazione è regolata in parte dal modo in cui udiamo la nostra voce cosicché la bambina *B* avrebbe potuto essere sorda. A ciascuna bambina furono perciò presentati stimoli acustici e vennero registrate le onde cerebrali evocate in risposta. Quando i suoni furono presentati alla sola bambina *A*, ottenemmo potenziali evocati dalla sua parte della testa e nessuna risposta dalla parte della bambina *B*. Quando presentammo i suoni alla bambina *B* non ottenemmo alcun potenziale evocato né dalla sua parte della testa né da quella della bambina *A*. Questa assenza di risposta avrebbe potuto esser dovuta a un danno nella capacità uditiva della bambina *B*, ma alla luce di altri elementi pensammo che potesse essere attribuita alle interpretazioni altamente variabili e oscillanti date dalla bambina delle onde cerebrali di fondo, che mascheravano probabilmente i mutamenti piú sottili che hanno luogo in risposta a stimoli acustici. Quando usammo lampi di luce per verificare le risposte delle due gemelle a stimoli visivi, trovammo che l'attività cerebrale di ciascuna delle due bambine era inequivocabilmente autonoma.

Quando le due gemelle ebbero raggiunto l'età di sei mesi, le sottoponemmo nuovamente a un test e trovammo

che c'era stato un mutamento inatteso. Le grida delle due bambine apparivano ora esattamente uguali. Ogni volta che la bambina *A* gridava, la bambina *B* si aggregava immediatamente e le due gemelle vocalizzavano con notevole omofonia. Divenne cosí estremamente difficile ottenere registrazioni di grida individuali. Quando riuscimmo a ottenere spettrogrammi acustici di singole grida, essi non presentavano differenze significative. Le grida anormali, discontinue della bambina *B* erano scomparse. Poco dopo fu compiuto il tentativo di separare le gemelle chirurgicamente, ma entrambe morirono nel corso dell'operazione.

In che modo bambine che avevano in precedenza « personalità vocali » distinte avevano potuto sviluppare grida identiche in soli tre mesi? La spiegazione da noi suggerita è che le due gemelle si influenzarono reciprocamente col loro comportamento vocale. Poiché la bambina *A* aveva un grido forte, normale, è probabile che la bambina *B* imparasse o almeno acquisisse in qualche modo il suo grido normale dalla gemella a lei unita. Da osservazioni di altri gemelli è noto che in certe circostanze i gemelli possono esercitare reciprocamente un'influenza cosí forte da sviluppare una sorta di linguaggio segreto inteso solo da loro.

Il pianto è tra i primi atti sociali di un bambino. Esso segna il passaggio da parte del bambino da uno stato di dipendenza muta dalla madre a una capacità di comunicare col mondo. Benché siano stati riferiti casi di vocalizzazione fetale *in utero* (noi non ne abbiamo mai scoperti), solitamente si considerano l'inizio della produzione di suoni umani i primi atti respiratori consecutivi alla nascita. James F. Bosma, del National Institute of Dental Research, ha dimostrato che i suoni alla nascita sono normalmente combinazioni di respiro affannoso, di tosse e di sforzi per respirare, i quali coincidono con la lotta compiuta dal neonato per adattarsi al suo nuovo ambiente. Altri studiosi hanno dimostrato che il primo comportamento vocale dei neonati risente del tipo di farmaci anestetici somministrati alla madre durante le doglie e della prontezza con cui il cordone ombelicale viene legato dopo la nascita.

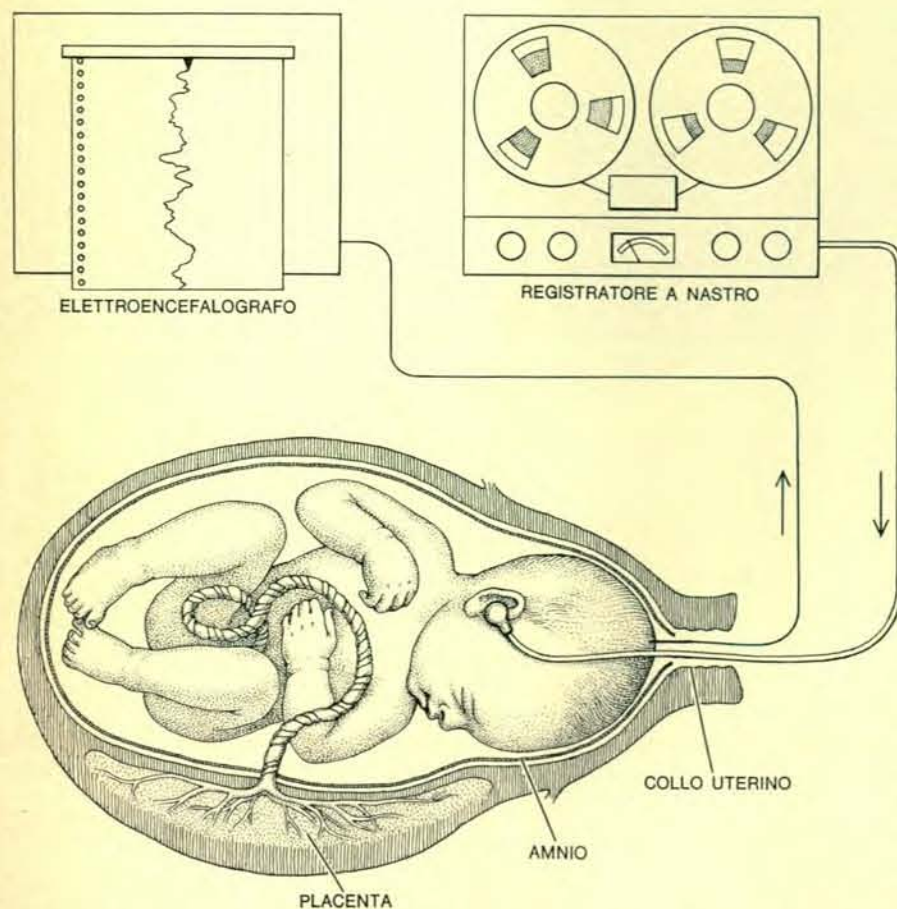
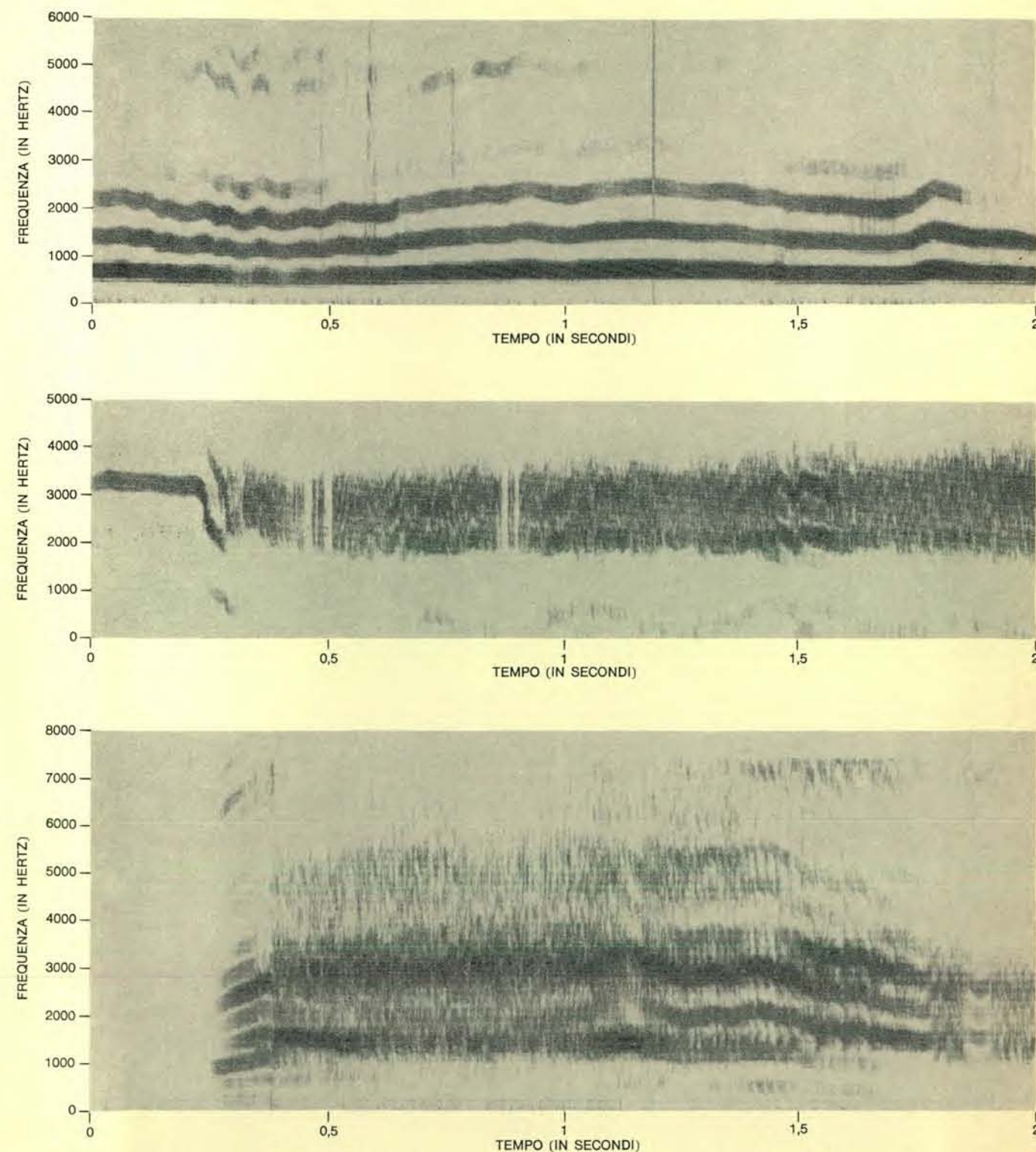
Uno studio attento del comportamento uditivo e vocale del neonato può rivelare secondo noi in modo preciso come si sviluppi la comunicazione fra il bambino e le persone del suo ambiente. Il problema consiste nel trovare aspetti del comportamento infantile che possano essere misurati oggettivamente e poi stabilire relazioni fra queste misurazioni e processi di sviluppo nel bambino da un lato ed eventi nell'ambiente dall'altro. In taluni casi tentativi di trovare relazioni del genere sono stati coronati da successo. In varie malattie causate da aberrazioni cromosomiche i bambini presentano grida caratteristiche. In un'infermità avente come conseguenza una laringe piccola e debole, il bambino ha un vagito acuto cosí caratteristico che la malattia è chiamata sindrome *cri du chat* (miagolio di gatto).

Recentemente siamo passati allo studio dei processi cerebrali associati alle prestazioni uditive del bambino. In una ricerca abbiamo tentato di stabilire se un feto avrebbe risposto a suoni nell'ultimo trimestre della gravidanza. Molte madri hanno osservato mutamenti nel comportamento di un feto in seguito a un rumore forte. Numerosi psicologi hanno tentato di condizionare il feto a rispondere a suoni, ma non è chiaro se le risposte osservate siano il risultato di un condizionamento fetale o del comportamento della madre.

È probabile che un feto, in virtú dello stadio avanzato di sviluppo del suo sistema uditivo, abbia un qualche genere di capacità uditiva. Al quinto mese di gestazione, le strutture dell'orecchio medio e dell'orecchio interno hanno già raggiunto le dimensioni di quelle dell'adulto. La capacità di questo sistema di ascolto prenatale nell'elaborare i suoni non è stata però ancora dimostrata in modo convincente.

Nel nostro studio sei donne gravide si prestarono volontariamente a questo esperimento all'inizio delle doglie. Un auricolare in miniatura fu introdotto nella cavità uterina attraverso il collo dell'utero già parzialmente dilatato e fu diretto lungo la testa del feto fino a una regione in prossimità dell'orecchio. Un elettrodo fu poi delicatamente applicato al cranio del feto e i fili dell'elettrodo furono collegati con l'apparecchiatura per la registrazione delle onde cerebrali (*si veda la figura a lato*).

Gli stimoli sonori cui il feto fu sottoposto constavano di scoppi di un suono di 1000 hertz a un livello moderatamente alto. Trovammo che soltanto un feto presentava una risposta coerente al suono. È degno di nota il fatto che la madre del feto che aveva risposto agli stimoli era l'unica a cui non fosse stato somministrato un farmaco per stimolare le contrazioni dell'utero.

Il controllo elettronico del feto durante le doglie è stato introdotto da Edward H.G. Hon, dell'Istituto di medicina della University of Southern California. Egli ha dimostrato che le registrazioni del ritmo cardiaco del feto e della pressione intrauterina forniscono informazioni clinicamente preziose sulla condizione del feto. Nel nostro lavoro non ci siamo limitati a esaminare le risposte dei feti a stimoli acustici, ma abbiamo controllato anche

La figura illustra in forma schematica le apparecchiature per il controllo delle risposte del feto durante le doglie. Un elettrodo flessibile in acciaio inossidabile è introdotto nella cavità uterina attraverso il collo dell'utero, che ha cominciato a dilatarsi in seguito all'inizio delle doglie. L'elettrodo è applicato alla testa e segnala le variazioni nell'attività delle onde cerebrali. Attraverso il collo dell'utero è introdotto anche un auricolare in miniatura, che viene fatto scivolare fino a farlo giungere in prossimità dell'orecchio. In una ricerca condotta su sei gestanti prestatesi volontariamente, soltanto un feto diede una risposta coerente ai suoni inviatigli attraverso l'auricolare.

Questi spettrogrammi acustici illustrano il pianto anormale di bambini sofferenti. Un'anomalia cromosomica nota come sindrome *cri du chat* (miagolio) dà origine a un vagito tipico (*in alto*). Una bambina affetta da ipoglicemia emette un grido privo quasi totalmente di componenti a bassa frequenza (*in centro*). I suoni emessi da un bambino che piangeva quasi costantemente avevano un tono molto alto ed erano strozzati (*in basso*). Il bambino morí per cause imprecisate all'età di otto mesi.

il ritmo cardiaco del feto, le sue onde cerebrali e i tipi di contrazione uterina; il tecnico alle apparecchiature era Phillip J. Goldstein del San Francisco General Hospital. Un calcolatore riceve i vari segnali e calcola l'ampiezza integrata dell'elettroencefalogramma fetale, il ritmo cardiaco medio del feto e la pressione intrauterina media per ogni periodo di cinque secondi. I valori sono stampati da una telescrivente. Il computer produce anche un'immagine grafica nella quale si vede in quale relazione stiano fra loro l'encefalogramma e il ritmo cardiaco del feto e la pressione intrauterina. Questi metodi forniscono all'ostetrico un'immagine numerica e bidimensionale precisa e relativamente obiettiva della funzione cerebrale e cardiovascolare del feto durante le doglie. Vengono cosí prontamente in evidenza eventuali situazioni suscettibili di influire negativamente sul neonato durante le doglie e arrecargli per esempio danni cerebrali.

Uno sviluppo recente in questo campo è costituito da un sistema ottico che utilizza un fascio laser, producendo su uno schermo televisivo un'immagine grafica delle onde cerebrali del feto pochi istanti dopo che esse sono state registrate. Con questo sistema abbiamo potuto dimostrare che la mepivacaina, un farmaco usato spesso durante le doglie per alleviare il dolore della gestante, riduce l'ampiezza delle onde cerebrali del feto. Abbiamo sviluppato inoltre un metodo per la trasmissione, durante il rilevamento stesso, di dati fisiologici di un feto da un ospedale all'altro su normali linee telefoniche. Questa possibilità è stata dimostrata in un collegamento tra il Mercy Hospital a Redding, California, e il San Francisco General Hospital, che si trovano a circa 300 chilometri l'uno dall'altro. Il sistema rende possibile un consulto e un'azione istantanei nel caso di infermità del feto.

In taluni casi una settimana o due dopo la nascita non sono ancora individuabili i segni neurologici di danni cerebrali, cosicché riesce difficile l'identificazione dei bambini che non si svilupperanno normalmente. Se il controllo elettronico esercitato durante le doglie avrà dato adito al sospetto di qualche disturbo generalizzato, forse un qualche tipo di esame di routine potrà rivelare sintomi piú specifici di danno cerebrale. Come abbiamo visto, l'analisi spettrografica del pianto del neonato e la misurazione di potenziali elettrici evocati da stimoli uditivi e visivi forniranno una base quantitativa molto promettente ai fini di un piú accurato controllo clinico di alcuni neonati.

GIOCHI MATEMATICI

di Martin Gardner

I "bastoncini" di Nepero, l'eccentrico padre dei logaritmi

Nel suo famoso *Budget of Paradoxes*, Augustus De Morgan definiva un «*graphomath*» una persona ignorante di matematica che cerchi di descrivere un matematico. Per darne un esempio, egli si riferiva al secondo capitolo del romanzo di Sir Walter Scott *The Fortunes of Nigel*, in cui David Ramsay, un bizzarro orologiaio e matematico dilettante, chiama a testimoni «le ossa dell'immortale Nepero!»

Da questo brano è difficile dire se Scott si esprimesse cosí per ignoranza o se intendesse invece attribuire a Ramsay un'osservazione ignorante o scherzosa. In ogni caso i «*Napier's bones*» non hanno nulla a che fare con i resti mortali del barone John Napier (1550-1617), volgarmente detto Nepero, il matematico scozzese che scoprí i logaritmi e che fu il primo matematico importante della Gran Bretagna. L'espressione si riferisce a una serie di bastoncini numerati inventati da Nepero per facilitare la moltiplicazione. Esamineremo fra poco il suo metodo; vediamo ora alcune notizie sul personaggio.

Il padre, Sir Archibald Napier, direttore della zecca scozzese, aveva esattamente 16 anni quando John nacque. John entrò tredicenne all'università di St. Andrews, la lasciò senza aver conseguito un titolo di studio, ereditò il castello e le proprietà di famiglia a Merchiston (che oggi fa parte di Edimburgo), si sposò ed ebbe un figlio e una figlia, rimase vedovo, si risposò e conservò la simmetria con cinque figli e cinque figlie. La Riforma protestante aveva avuto inizio in Scozia pressappoco nel periodo in cui egli era nato, e negli anni trascorsi all'università di St. Andrews egli divenne un protestante appassionato, con un'inclinazione irresistibile a spiegare le profezie bibliche. Nel 1593 pubblicò quello che egli considerò sempre il suo capolavoro (molto piú importante dei logaritmi), ossia la *Plaine discovery of the Whole Revelation of Saint John*, il cui titolo completo suona pressappoco: «Spiegazione chiarificatrice dell'intera Apocalisse di San Giovanni, esposta in due trattati, nel primo dei quali se ne ricerca e dimostra la vera interpretazione, mentre nel secondo la si applica parafrasticamente e storicamente al testo, per opera di John Napier di Merchiston, junior. In appendice vengono riprodotti taluni Oracoli della Sibilla, in accordo con l'Apocalisse e con altri luoghi della Scrittura. Edimburgo, per cura di Robert Walde-grave, stampatore di Sua Maestà Reale, 1593. Con privilegio reale.»

Fu questa la prima grande opera scritta in Scozia sulla Bibbia e uno fra i tentativi piú approfonditi mai fatti prima o dopo d'allora per esplorare il simbolismo dell'Apocalisse. È strano che oggi, quando molti studenti sembrano piú interessati al Secondo Avvento che alla attualità politica, non ci sia una ristampa disponibile del trattato di Napier. Esso esercitò una grandissima influenza ai suoi tempi, ed ebbe 21 edizioni in inglese e traduzioni in numerose lingue europee.

Forse la ragione principale dell'indisponibilità del libro è da ricercarsi nel fatto che Napier compí un leggero errore di calcolo nella previsione della fine del mondo. Egli era stato fortemente influenzato dalle speculazioni religiose di Michele Stifel, un algebrista tedesco il quale aveva dimostrato che il papa Leone X era l'anticristo ridisponendo le cifre romane contenute in LEO DECIMVS a formare DCLXVI, o 666, il celebre «segno della bestia.» Ma da dove Stifel trasse la X? Da Leo X e dal fatto che LEO DECIMVS ha 10 lettere. E che cosa ne fu della M? Stifel la lasciò fuori perché stava per *mysterium*. Stifel predisse che il mondo sarebbe finito il 3 ottobre del 1533. Napier, nato dopo quella data, aveva la certezza di fatto che c'era stato un errore. Egli stabilí che il vero anticristo era il papa del 1593. Dio aveva prescritto che tra la creazione del mondo e la sua distruzione trascorressero esattamente 6000 anni. Sussistendo qualche incertezza circa la data esatta della creazione, Nepero fissò la fine del mondo al periodo compreso tra il 1688 e il 1700.

Nepero inizia il libro chiedendo scusa per averlo scritto in una lingua cosí indegna come l'inglese, e lo conclude rivolgendosi al papa nel modo seguente: «In conclusione, se tu, *Roma*, dichiari di essere riformata e di credere nel vero cristianesimo, allora credi in San *Giovanni* il Discepolo, colui che Cristo amava, e che qui, nella sua Apocalisse, proclama la tua rovina, ma se tu rimani pagana nei tuoi pensieri piú riposti, credendo negli Oracoli antichi delle *Sibille*, considerati un tempo con reverenza nel tuo *Campidoglio*, allora anche questa *Sibilla* proclama la tua rovina. Pentiti perciò in questo tuo ultimo respiro, se ti sta a cuore la tua Eterna salvezza. *Amen.*»

«È strano – commentava De Morgan nel suo *Budget* – che Nepero non si rendesse conto del fatto che questo suo appello non poteva aver successo, a meno che le profezie dell'Apocalisse fossero erronee.»

Dopo aver messo in chiaro i misteri apocalittici, Nepero rivolse la sua ingegnosità alla ricerca di modi per difendere la Scozia contro la minaccia di un'invasione da parte della Spagna cattolica. Il suo scritto del 1596, *Secrett Inventionis*... («Invenzioni segrete, utili e necessarie in questi tempi per la difesa di quest'isola e per opporsi agli stranieri, nemici della verità divina e della religione»), descrive tre invenzioni: specchi per incendiare le navi nemiche (gli specchi ustori di Archimede!), una mitragliatrice e una carrozza metallica (cioè un carro armato), dall'interno della quale i soldati potevano far fuoco attraverso feritoie laterali.

Il successivo libro di Napier, il cui titolo latino comincia con *Mirifici logarithmorum canonis descriptio*... (Descrizione della meravigliosa regola dei logaritmi), apparve nel 1614. Fu questo il libro in cui Nepero spiegò i logaritmi, in cui assegnò loro questo nome (il vocabolo fu coniato da lui) e in cui diede al mondo la sua prima tavola logaritmica. È stato spesso sottolineato che se fossero allora stati d'uso comune gli esponenti delle potenze, i logaritmi sarebbero stati riconosciuti immediatamente come un grande ausilio, ma Nepero li concepí senza alcun riferimento ad esponenti. I logaritmi di Nepero si fondavano su una base scomoda, di poco minore di 1. (Non si trattava cioè dei logaritmi naturali con base *e*, anche se questi ultimi divennero noti come logaritmi neperiani.) Il geometra londinese Henry Briggs si rese immediatamente conto che 10 era la base piú conveniente per calcoli logaritmici nel sistema decimale, e Nepero fu subito d'accordo.

Naviganti e astronomi – tra questi particolarmente Keplero – trovarono i logaritmi di base 10 di utilità inestimabile, e Briggs e altri dedicarono anni e anni di fatica improba a preparare tavole logaritmiche sempre migliori. (Oggi si fa prima a calcolare un logaritmo su una calcolatrice elettronica tascabile – che impiega solo un microsecondo – che non a cercarlo su un libro!) Nell'opera postuma *Mirifici logarithmorum canonis constructio*... (1619), John Napier spiegò in che modo avesse calcolato i suoi logaritmi originali. In quest'impresa, per la prima volta nella storia, egli aveva fatto un uso sistematico del punto decimale, collocandolo al di sopra della linea di base, esattamente com'è usato ancor oggi in Inghilterra. [Com'è noto, nei paesi di lingua inglese si usa il punto, laddove noi usiamo invece la virgola, per dividere i numeri interi dai decimali, mentre vi si fa uso della virgola (invece del nostro punto o della semplice spaziatura) per dividere i numeri superiori al migliaio in gruppi di tre cifre.]

Tra i molti aneddoti su Nepero, due particolarmente spassosi sono narrati da Howard W. Eves nella sua divertente opera *In Mathematical Circles*. I piccioni di un vicino volavano nella proprietà di Nepero e si cibavano del suo grano; Nepero disse allora al vicino che avrebbe preso i piccioni a titolo di risarcimento. Il vicino rispose che era libero di impadronirsi di ogni piccione che fosse riuscito a catturare vivo. Nepero sparse nella sua proprietà piselli imbevuti di acquavite e i piccioni si trovarono ben presto in uno stato di tale torpore che egli non ebbe alcuna difficoltà a catturarli e infilarli tutti in un sacco.

Era quella un'epoca in cui quasi tutti in Scozia (compreso Nepero) credevano nell'astrologia e nella magia nera. Un giorno Nepero riuní i suoi servitori e disse loro che il suo gallo nero aveva il potere occulto di dirgli quale servitore avesse compiuto furti nella proprietà. I servitori furono invitati uno per volta a entrare in una stanza debolmente illuminata e accarezzare il dorso del gallo. Come Nepero aveva previsto, soltanto la persona colpevole, temendo di esporsi, non lo avrebbe fatto. Nepero aveva coperto le penne nere del

4896 × 7 = 34 272

gallo di fuligine, e cosí soltanto il servitore colpevole venne fuori con le mani pulite.

A quell'epoca c'era anche grande interesse per il calcolo. Le persone comuni calcolavano con l'aiuto delle dita, ma c'erano anche abili matematici che si divertivano a eseguire calcoli lunghi e tediosi. Il pallino di Nepero era quello di trovare procedimenti che consentissero di semplificare tale lavoro. I logaritmi furono, ovviamente, la sua invenzione migliore, ma nel 1617 (l'anno della sua morte) egli pubblicò un libricino, intitolato *Rabdologia*, che spiegava altri tre metodi di calcolo. Il titolo del libro era anche il nome del primo metodo, quello dei «bastoncini di Nepero», detto anche, in Inghilterra, dei «*Napier's bones*» perché egli usava bastoncini che erano fatti spesso di ossa di animali.

Il lettore è invitato a costruirsi una serie di bastoncini di Nepero usando 11 strisce di cartone pesante (o bastoncini di ghiaccioli o altre strisce di legno) e riportando su di esse i valori indicati nella figura della pagina a fronte. Il bastoncino indice non è essenziale ma facilita l'individuazione della riga desiderata. Ciascun bastoncino ha in alto un numero; sotto di esso, dall'alto in basso, sono i prodotti della moltiplicazione di tale numero per i numeri da 1 a 9. Chiaramente, l'insieme dei bastoncini non è niente di piú di una tavola pitagorica tagliata a strisce in modo da poter essere manipolata manualmente, con una striscia dello zero (o striscia nulla) aggiunta come «segnaposto».

Il procedimento è semplicissimo. Supponiamo che si desideri moltiplicare

BASTONCINO INDICE

Rabdologia: i «bastoncini di Nepero».

4896 per 7. Si collocano l'uno accanto all'altro i bastoncini del 4, dell'8, del 9 e del 6, col bastoncino indice sulla sinistra (*si veda la figura della pagina precedente*). Consideriamo soltanto la riga 7 (il moltiplicatore). Si scriva 2, l'ultima cifra della riga, come cifra finale del prodotto. La cifra successiva, verso sinistra, si ottiene sommando la coppia di cifre successive, procedendo verso sinistra sempre sulla riga 7 (in questo caso il 4 della colonna del 6 e il 3 della colonna del 9). Poniamo cosí il 7 come penultima cifra del prodotto. La somma della coppia successiva (6 + 6 = 12) è superiore a 9: scriviamo perciò 2 come terz'ultima cifra del prodotto e riportiamo 1. La coppia successiva, 5 e 8, sommata dà 13, che con l'1 di riporto fa 14. Scriviamo 4 e riportiamo di nuovo 1. L'ultima cifra della riga è 2. 2 + 1 fa 3, cosicché 3 sarà la cifra iniziale del nostro prodotto. Abbiamo ottenuto cosí la risposta esatta, 34 272, usando soltanto l'addizione. Ovviamente, conoscendo la tavola pitagorica si può procedere esattamente come con i bastoncini. Ai tempi di Nepero la capacità di calcolo era però molto bassa e i bastoncini ottennero perciò un successo immediato.

Per moltiplicare 4896 per un numero maggiore, diciamo 327, è necessario ottenere tre prodotti parziali, e sommarli, nel modo usuale. In altri termini, si scriva 37 272 (il prodotto di 4896 per 7), quindi si riportino sotto i prodotti ottenuti dalle righe 2 e 3, sfalsandoli di un posto da sinistra nel modo solito,

34272
9792
14 688

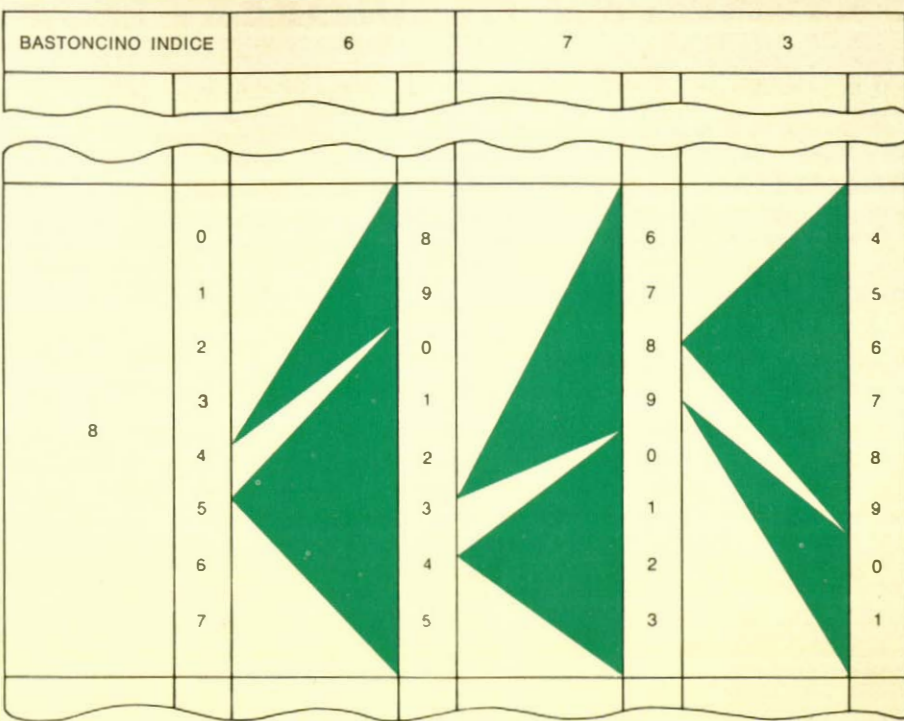
e sommando si ottiene poi il prodotto finale.

I bastoncini sono di scarsa utilità a meno che non se ne possegga piú di una serie perché un moltiplicando può contenere due volte la stessa cifra. I bastoncini di Nepero avevano la sezione quadrata e ciascuna faccia di un bastoncino corrispondeva a una striscia della nostra serie di cartone. Nepero dispose le quattro colonne in modo che le cifre di testa delle facce opposte di ogni bastoncino dessero come somma 9. Ecco le quadruple della serie di 10 bastoncini di Nepero:

0, 1, 9, 8	1, 3, 8, 6
0, 2, 9, 7	1, 4, 8, 5
0, 3, 9, 6	2, 3, 7, 6
0, 4, 9, 5	2, 4, 7, 5
1, 2, 8, 7	3, 4, 6, 5

È chiaro che una tale serie di 10 bastoncini può trattare tutti i moltiplicandi di 10 cifre o meno che si possono formare con i bastoncini stessi, ma molti moltiplicandi non possono essere formati, cosicché era consigliabile possederne piú di una serie. Come piccolo problema combinatorio, il lettore provi a determinare qual è il massimo moltiplicando che si può formare con una serie di bastoncini di Nepero, tale che, con la stessa serie, si possano formare anche tutti i moltiplicandi di esso minori. Come secondo esercizio, si trovi il corrispondente moltiplicando massimo per due serie di bastoncini di Napier. (Entrambe le risposte verranno date il mese prossimo.)

I bastoncini di Nepero possono essere usati anche per la divisione, ma il procedimento è piú complicato e non val la pena di usarlo. Nelle divisioni brevi si devono scegliere i bastoncini che formino il dividendo desiderato sulla riga corrispondente alla cifra del divisore, e poi si leggerà il quoziente in alto. Se il dividendo non può essere formato, si formi il numero piú grande possibile che sia inferiore al dividendo, poi si sottragga quel numero dal dividendo, ottenendo cosí il resto. In divisioni lunghe, i bastoncini possono essere usati per determinare i successivi prodotti del divisore e ciascuna cifra del quoziente.

Il fascino dei bastoncini di Nepero risiede nella loro semplicità. Se però siamo disposti a complicarli un pochino, elimineremo la seccatura di dover tenere a mente le unità da riportare. Il modo piú ingegnoso per eliminare questo fastidio fu inventato attorno al 1890 da Henry Genaille, un ingegnere civile francese. La figura che illustra questi bastoncini non ha quasi bisogno di spiegazione (*si veda la figura nella pagina a fronte*). Essi funzionano esattamente come quelli di Nepero, tranne che per il fatto che il prodotto si legge direttamente da sinistra a destra. Si cominci dalla cifra che si trova a destra in alto della riga desiderata. La cifra successiva è quella indicata dal vertice del triangolo colorato (a sinistra della cifra precedente). Da questo punto in avanti si passa da ciascuna cifra al triangolo in colore che si trova direttamente alla sua sinistra e si va alla cifra indicata dal suo vertice. Per esempio, per moltiplicare 673 per 8 si comincia col 4 a destra in alto (*si veda la figura in questa pagina*) e si osservi con quanta facilità si possa muovere verso sinistra attraverso la catena di triangoli per ottenere il prodotto 5384.

I bastoncini calcolatori di Henry Genaille.

673 × 8 = 5384

Sia i bastoncini di Nepero sia quelli di Genaille sono congegni tecnici meravigliosi perché non è difficile comprenderne il meccanismo e ciò consente di capire piú a fondo il procedimento della moltiplicazione. Se il lettore prova qualche difficoltà nel capire il funzionamento dei bastoncini di Genaille, può trovarne la spiegazione in *Genaille's Rods: An Ingenious Improvement on Napier's* di B.R. Jones, in «The Mathematical Gazette», vol. 48, febbraio 1964, pagine 17-22, l'articolo da cui sono state tratte le nostre illustrazioni.

Il secondo metodo di calcolo illustrato nella *Rabdologia* si fonda sulla disposizione di lamine metalliche all'interno di una scatola ed è troppo complicato e troppo poco pratico per spiegarlo qui. Il terzo metodo di Nepero, che egli considerava soprattutto un divertimento, richiede soltanto una scacchiera e una certa quantità di gettoni. Muovendo i gettoni come si muovono le torri e gli alfieri, si possono eseguire somma, sottrazione, moltiplicazione e divisione, il tutto nel sistema binario.

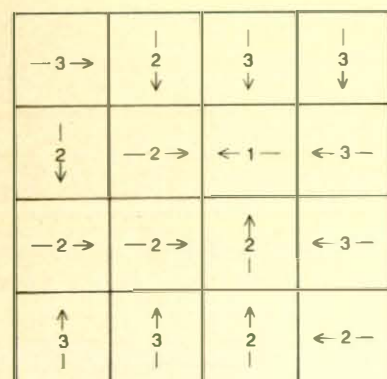
Il mese scorso presentai una raccolta di problemi di vario genere. Ecco le soluzioni:

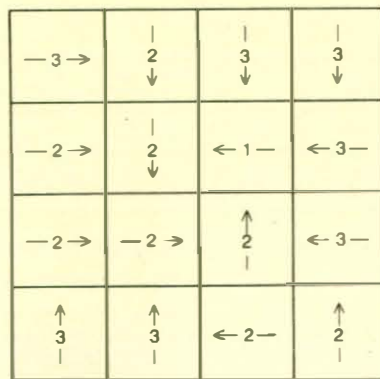
1. I tre modi di formare percorsi di lunghezza massima sul campo 4 per 4 sono illustrati nella figura in basso della pagina seguente.

2. Delle cinque coppie composte da marito e moglie nessuna strinse piú di otto mani. Quindi, se nove persone stringono ognuna un numero differente di mani, i numeri devono essere: zero, uno, due, tre, quattro, cinque, sei, sette e otto. La persona che strinse otto mani doveva essere sposata con chi non ne strinse neanche una (altrimenti avrebbe potuto stringerne solo sette). Analogamente, la persona che strinse sette mani doveva essere sposata con la persona che strinse solo una mano

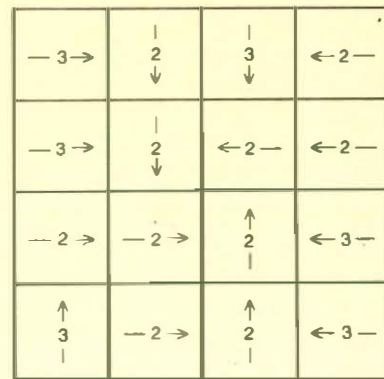
Soluzione del problema delle strette di mano.

(la mano, cioè, della persona che strinse otto mani). La persona che strinse sei mani doveva essere sposata con quella che ne strinse due e la persona che ne strinse cinque doveva essere sposata con chi ne strinse tre. La sola persona rimasta, che strinse la mano a quattro persone, è mia moglie (*si veda la figura in alto*).

3. Un poligono convesso di 11 lati può essere formato, a partire da quadrati e triangoli equilateri di lato unitario, nel modo indicato nella figura in alto della pagina seguente. Gli angoli possibili per un poligono formato a partire da questi elementi son di 60, 90, 120 e 150 gradi. Per ottenere il poligono col maggior numero possibile di lati, tutti gli angoli dovranno essere di 150 gradi e il numero dei lati sarà 12.

4. Solo il nono enunciato è vero.

5. Una soluzione per costruire una recinzione di pentamini, di forma arbitraria, per una superficie dell'area di 128 unità quadrate, è indicata nella seconda figura dall'alto della pagina seguente. Non è l'unica soluzione possibile.

6. Un tavolo quadrato può essere sempre collocato in modo che, anche su un pavimento ondulato, tutte le gambe tocchino terra. Per dimostrare questa proposizione collochiamo il tavolo in un posto qualsiasi. Supponiamo che solo tre gambe, *A,B,C* tocchino terra e che la gamba *D* resti sospesa (*si veda la terza figura dall'alto della pagina seguente*). È sempre possibile che tre gambe tocchino terra dato che, presi comunque tre punti nello spazio, essi segnano sempre gli angoli di un triangolo. Ruotate ora il tavolo di 90 gradi intorno al centro, mantenendo le gambe *A* e *B* sempre attaccate al pavimento. Il tavolo si troverà allora in una posizione in cui *C* è la sola gamba a non toccare terra.

Durante la rotazione *D* è venuta a contatto col pavimento, mentre *C* se ne è staccata. Ma *D* deve aver toccato il pavimento prima che *C* se ne staccasse, altrimenti vi dovrebbe essere una posizione in cui solo *A* e *B* toccano terra, mentre sappiamo che è sempre possibile far toccare terra a tre gambe. In qualche momento della rotazione, dunque, tutte e quattro le gambe devono essere a contatto col suolo. Un ragionamento analogo può essere applicato a tavoli vacillanti rettangolari: basta farli ruotare di 180 gradi.

7. Per ottenere una pieghettatura formante un disegno a mosaico, con tasselli esagonali regolari, su di un piccolo foglio di carta, arrotolatelo, per prima cosa, fino a formare un tubo di circa un centimetro di diametro. Con l'indice e il pollice della sinistra prendete una estremità del tubo, schiacciandola fino ad appiattirla. Mantenendo la presa con la sinistra, stringete ora, con l'indice e il pollice della mano destra, il tubo più vicino che potete alla zona già compressa dalla sinistra, facendo una piega ad angolo retto rispetto alla prima. Stringete con forza con entrambe le mani e nello stesso tempo spingete le due pieghe una accanto all'altra, strettamente, in modo che la piegatura risulti il più possibile netta. Ora, mentre la mano destra tiene la sua piega, la sinistra fa una terza piegatura adiacente e perpendicolare alla seconda. Continuare così, alternando le mani man mano che si fa passare il tubo, finché tutto il tubo è stato pieghettato. (I bambini spesso fanno un gioco analogo con le cannucce delle bibite per fare delle « catene ».) Svolgete il foglio. Lo troverete ricoperto da un disegno a pieghe esagonali in modo assolutamente incomprensibile per il non iniziato.

—3→	2↓	3↓	3↓
2↓	—2→	←1—	←3—
—2→	—2→	↑2	←3—
↑3	↑3	↑2	←2—

—3→	2↓	3↓	3↓
—2→	2↓	←1—	←3—
—2→	—2→	↑2	←3—
↑3	↑3	←2—	↑2

—3→	2↓	3↓	←2—
—3→	2↓	←2—	←2—
—2→	—2→	↑2	←3—
↑3	—2→	↑2	←3—

Risposta ai giri di frecce.

8. Per ottenere la posizione di scacchi fornita il mese scorso, disponete i pezzi come indica la figura in basso e fate le mosse seguenti:

Bianco	*Nero*
1	B-Q4 (scacco)
2 P-B4	P prende P *en passant* (doppio scacco)
3 K prende P(scacco)	

Levando il re bianco otterremo la posizione desiderata.

9. La chiave per semplificare le tre equazioni sulle polipotenze è la legge fondamentale

$$(a^b)^c = a^{(b \times c)}.$$

Applicando questa legge alla prima equazione,

$$x^{(\overleftarrow{x^{x^{x}}})} = (x^{x})^{\overrightarrow{x}}$$

otteniamo

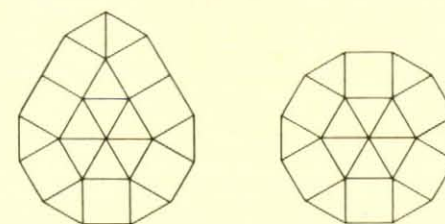
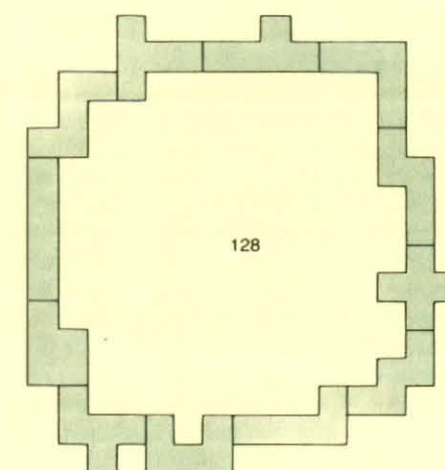
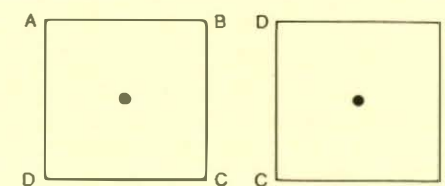
$$x^{(\overleftarrow{x^{x^{x}}})} = x^{(\overrightarrow{x^{2}})}$$

Le due x alla base sono uguali, quindi i loro esponenti racchiusi tra parentesi sono uguali. Concelliamo le x alla base e ripetiamo la procedura:

$$x^{(\overleftarrow{x^{x}})} = x^{2}$$

Cancelliamo di nuovo le x alla base ottenendo $x^x = 2$, che dà alla x il valore di 1,55961+.

La stessa procedura semplifica la seconda equazione (quella in cui si eguagliano due tetrapotenze, una calcolata verso l'alto ed una verso il basso) ottenendo $x^x = 3$ ed $x = 1.82545+$. Ogni equazione successiva accresce di uno il valore di x^x. La terza equazione si riduce a $x^x = 4$, cioè $x = 2$.

La procedura generale è quella di rimpiazzare una scala calcolata verso l'alto con un numero pari al numero degli x occorrenti in essa meno uno, e rimpiazzare la scala calcolata verso il basso con una avente due x di meno. (Esempio: verso l'alto–5 = verso il basso–5 si riduce a 4 = verso il basso–3.)

Poligoni convessi di 11 e 12 lati.

Il più grosso « recinto » di pentamini.

La dimostrazione del tavolo vacillante.

Analisi retrograda.